REPORT RESUMES

ED 003 076 24

A COMPARISON OF THE EMPIRICAL VALIDITY OF SIX TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES.

BY- MUELLER, MAX W.
GEORGE PEABODY COLL. FOR TEACHERS, NASHVILLE, TENN.

REPORT NUMBER CRP-S-059 PUB DATE AUG 65
CONTRACT OEC-4-10-199
EDRS PRICE MF-$0.27 HC-$5.56 139P.

DESCRIPTORS- *EDUCABLE MENTALLY HANDICAPPED, *TEST VALIDITY, *COMPARATIVE ANALYSIS, *DIAGNOSTIC TESTS (EDUCATION), INTELLIGENCE TESTS, ACHIEVEMENT TESTS, PREDICTIVE ABILITY (TESTING), *MEASUREMENT INSTRUMENTS, PRIMARY MENTAL ABILITIES TEST (PMAT), COLOURED PROGRESSIVE MATRICES (CPM), STANFORD BINET INTELLIGENCE SCALE, PEABODY PICTURE VOCABULARY TEST (PPVT), PICTORIAL TEST OF INTELLIGENCE (PTI), ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES (ITPA), NASHVILLE, TENNESSEE

AN INVESTIGATION OF THE VALIDITY OF INTELLIGENCE AND OTHER TESTS USED IN THE DIAGNOSIS OF RETARDED CHILDREN WAS PERFORMED. EXPERIMENTAL SAMPLES CONSISTED OF 101 CHILDREN SELECTED FROM SPECIAL CLASSES FOR EDUCABLE MENTALLY RETARDED (EMR) WHOSE AGES RANGED FROM 6.9 TO 10 YEARS AND WHOSE IQ SCORES RANGED FROM 50 TO 80. THE TESTS EVALUATED WERE (1) STANFORD-BINET, (2) ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES (ITPA), (3) PEABODY PICTURE VOCABULARY TEST (PPVT), (4) PICTORIAL TEST OF INTELLIGENCE (PTI), (5) COLOURED PROGRESSIVE MATRICES (CPM), AND (6) PRIMARY MENTAL ABILITIES TEST (PMAT). THESE TESTS WERE ADMINISTERED TO THE GROUP WITH VARIOUS LEARNING ABILITY TASKS. A STATISTICAL ANALYSIS SHOWED THAT THE PMAT, STANFORD-BINET, PTI, AND THE ITPA WERE VALID, WHILE THE OTHER TWO WERE LESS VALID. THE USE OF A COMBINATION OF ONE GROUP TEST AND ONE INDIVIDUAL TEST WAS ALSO SUPPORTED BY THESE RESULTS, AND THE CPM AND THE PPVT ARE NOT ADEQUATE SUBSTITUTES FOR MORE COMPREHENSIVE INSTRUMENTS THAT PREDICT LEARNING ABILITY AT AN EARLY AGE. (RS)

ED003076

S-059

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

A COMPARISON OF THE EMPIRICAL VALIDITY OF SIX TESTS OF ABILITY WITH EDUCABLE MENTAL RETARDATES

by

Max W. Mueller

Master of Science

Kansas State Teachers College

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
in the
Department of Special Education
of the
Graduate School
George Peabody College for Teachers
August, 1965

Approved:

Major Professor: ______________________

Minor Professor: ______________________

Dean of the Graduate School: ______________________

ED 003076

ACKNOWLEDGMENTS

Principle support for this project was obtained from the Small Grants Program of the Cooperative Research Branch of the United States Office of Education, under contract number OE4-10-199. Additional support was obtained from the Institute on Mental Retardation and Intellectual Development, Peabody College, Nashville, Tennessee.

Acknowledgment is also due Dr. Lloyd Dunn, Dr. Julius Seeman, Dr. William Griffin, Dr. Richard Woodcock, Dr. Carl Haywood, Dr. Laird Heal, and Mr. Don Dickerson for their assistance and advice in the development of this project. The author is particularly grateful to Dr. Dunn and Dr. Seeman for their extensive help, including reviewing the manuscript of the report.

The Nashville-Davidson County Metropolitan Public Schools also deserve recognition for their extensive help in carrying out this project. Recognition is due the administrative personnel who reviewed and approved the proposal and aided the investigator in the selection of the experimental group, and especially the principals and teachers in the participating schools who were extremely cooperative in adjusting their schedules to allow time and space for the examination of subjects.

The author also wishes to acknowledge the assistance of Evelyn Cromwell, Nancy Haywood, Rebecca Alonso, Cornell Lane,

Nettie Byrd, Richard Spong, Tom McConnell, Gayle Alexander, Frances Thompson, and Frances Rogers, the examiners who collected the majority of the test data for this investigation.

Credit is also due James Heustess, Reba Lewis, Karen Youngblood, Frances Howick, Lorraine Berenson, Genelle Porter, Nancy Lee, and Mildred Harrington for their assistance with the compilation of data and with editing and typing of the manuscript.

TABLE OF CONTENTS

CHAPTER PAGE

ACKNOWLEDGMENTS ii

TABLE OF CONTENTS iv

LIST OF TABLES vi

I INTRODUCTION 1

Identification of Terms 2

Purpose 4

Background 4

II METHOD 6

Subjects 6

Instruments 7

Procedures 10

Analyses of Data 14

III RESULTS 19

Descriptive Data 19

Correlational Analyses 19

Multiple Regression Analysis 24

Factor Analysis 27

IV DISCUSSION 35

Descriptive Data 35

Correlational Analyses 40

Factor Analysis 51

V SUMMARY, CONCLUSIONS, AND IMPLICATIONS 59

TABLE OF CONTENTS (continued)

PAGE

REFERENCES 64

APPENDIXES 68

A INSTRUMENTS USED IN THE STUDY 69

B INTERCORRELATIONS 103

C RAW TEST DATA 105

BIBLIOGRAPHY 120

LIST OF TABLES

TABLE PAGE

1 Reliability Reported on Test Manuals 11

2 Validity Scores on Predictor Tests with Mental Retardates 12

3 Scores Submitted to Correlational Analysis . . . 15

4 Means and Standard Deviations on Experimental Variables 20

5 Correlations of Predictor Tests with Criterion Scores 22

6 Correlations of Learning Tasks with Achievement Scores 23

7 Intercorrelations of Predictor Tests 24

8 Results of Multiple Regression Analysis of Subtests Scores 26

9 Variables Included in Factor Analysis 28

10 Principal-Axis Factor Analysis--Raw Factor Solution 29

11 Factor Pattern--Biquartimin Rotation 31

12 Factor Structure--Biquartimin Rotation 32

13 Factor Structure--Graphical Rotation 33

14 Intercorrelations of Biquartimin Factors 34

CHAPTER I

INTRODUCTION

Virtually since the beginning of the testing movement around the turn of the century (Binet, 1905), one of the primary criteria for the placement of children in special schools and classes for the educable mentally retarded has been a low score on an intelligence test. Stated broadly, the function of such a test is to predict the learning ability of children so they may receive educational treatment appropriate to their level of ability and rate of development. The effectiveness of such placement and treatment depends largely upon the validity of the test in measuring the level at which the child is functioning and predicting the rate at which he can be expected to progress. Thus the need for investigation of the validity of intelligence tests and other tests used in the diagnosis of retarded children is clear. In addition to the basic problem of obtaining validity data on psychometric tests, it is particularly desirable to obtain such data on a population of mentally retarded children in their early school years. As the number of special programs for the educable mentally retarded has increased, there has been a growing tendency for children to be placed in these programs within the first two or three years of school. This trend, coupled with at least suggestive data that currently popular intelligence tests are least valid with this younger, slower group, indicates the need for investigation of test validity with this specific group. Another indication of the need for evaluation of test validity is the growing concern with the efficacy of special classes for educable retardates (Kirk, 1964). It seems reasonable to assume that the adequacy of special classes is dependent, at least in part, on the validity of the instruments used in pupil placement. Thus, if it can be demonstrated that a given psychometric instrument is superior in terms of predicting learning ability, it should subsequently be possible to improve placement procedures.

Identification of Terms

Six psychometric instruments were evaluated in this study. Throughout this report these instruments have been identified as predictor tests.

1. The Stanford-Binet Intelligence Scale (S-B, Terman & Merrill, 1960)

2. The Illinois Test of Psycholinguistic Abilities (ITPA, McCarthy & Kirk, 1961)

3. The Peabody Picture Vocabulary Test (PPVT, Dunn, 1959)

4. The Pictorial Test of Intelligence (PTI, French, 1963)

5. The Coloured Progressive Matrices (CPM, Raven, 1947)

6. The Primary Mental Abilities Test (PMAT, Thurstone & Thurstone, 1962)

The validity of these tests was evaluated in terms of their correlations with a variety of criteria including tests of academic achievement and tasks requiring the subject to learn within the test situation. Two achievement tests were utilized in this investigation:

1. The Wide Range Achievement Test (WRAT, Jastak, 1945)

2. The New York Achievement Tests (NYAT, Wrightstone, _et al._, 1959)

Four tasks requiring the subject to learn within the testing session were also included among the criteria:

1. Coding (Wechsler, 1949)

2. The Modified Prognostic Reading Test (MPRT, Woodcock, 1964)

3. A paired associate task (PAT)

4. A simple alternation task (SAT)

For the purposes of this investigation these measures have been designated as criterion tests. In addition the validity of the predictor tests was evaluated in terms of correlations with three composite scores: (1) Criterion I (C_I), the mean of scores on the achievement tests; (2) Criterion II (C_{II}), the mean of the scores on the learning tasks; and (3) Criterion III (C_{III}), the mean of C_I) and C_{II}). This overall composite score (C_{III}) has been defined as the "demonstrated learning score" for the purposes of this study.

Purpose

The purpose of this investigation was to evaluate and compare the empirical validity of the six predictor tests. More specifically, the first objective of the study was to evaluate the six predictor tests in terms of their correlations with the demonstrated learning score, and with each of the individual criterion scores. This included both determining whether correlations were significantly different than zero, and determining if significant differences existed among the obtained correlations. The validity of the predictor tests was further evaluated by examination of the intercorrelations among these tests. Finally, the relationships among the various tests used in the study were evaluated further through the use of multiple regression analysis and factor analysis.

Background

According to Anastasi (1961) the validity of intelligence tests is most commonly evaluated in terms of correlations with measures of academic achievement or with previously established tests of intellectual ability. Denny (1964) has suggested the use of learning tasks as a possible criterion measure for the validation of psychometric tests of ability. Since it utilizes all three of these alternatives, the present study should be superior to any single approach. A review of the literature yielded no reference to studies using such a procedure.

Though a more thorough analysis of available validity data is presented in connection with the description of each test in Appendix A, a general statement seems appropriate here. An extensive body of knowledge concerned with various earlier versions of the S-B, PMAT, and Progressive Matrices exists, but little relates to the specific editions investigated in this study, S-B (1960), PMAT (1962), and CPM (1947). Review of the literature on all editions of these tests still yields little information specifically related to a retarded population. Fairly extensive research data have been collected regarding the validity of the PPVT, much of it dealing with the retarded. The problem in this instance is that results of previous studies have been extremely inconsistent. The final two instruments, ITPA and PTI, are of such recent vintage that very limited information is yet available concerning their validity. This is especially true of the PTI which has been available only since August, 1964.

CHAPTER II

METHOD

Information concerning the experimental sample, instruments, and procedures of collection and analysis of data is presented in this chapter.

Subjects

Experimental subjects were selected from children who had been identified as educable mentally retarded and enrolled in public school special classes. Children were included only if their S-B IQ scores were between 50 and 80. To stay within the range of some of the tests being evaluated, the group was further restricted to children from seven to 10 years of age.

The experimental group was composed of 101 subjects remaining from a total of 141 subjects selected to participate in this study. Of the 40 subjects excluded, 17 were eliminated because they were not available by the time testing commenced, six were dropped because they scored above 80 on the S-B, three because they obtained scores below 50, two because they were uncooperative during the testing, and 12 because they did not complete the entire test battery. The remaining 101 subjects (38 girls and 63 boys) made up the experimental group. The mean chronological age of the experimental group was 9 years with a range from 6 years 11 months to 10 years 4 months and a standard deviation of 9.35 months. This group was drawn from 17 special classes in 14 schools in the Nashville-Davidson County Metropolitan Schools.

Instruments

Twelve tests were administered to each subject. These are divided into six predictor tests and six criterion tests (tests of demonstrated learning ability). Admittedly, the distinction between prediction and demonstration of learning ability, at least in terms of test scores, is somewhat arbitrary. Nevertheless, a relatively clear distinction can be made. The predictor tests are designed essentially to predict the rate at which children should be able to learn, and to indicate their present level of intellectual functioning, although certainly they do so by making inferences from measurements of what the subject has learned. These tests are classified as predictors in this study on the basis of the intent of the test authors and of general identification as such among educators and psychologists.

The choice of predictor tests to be evaluated was largely an arbitrary decision based on the interests and experiences of the author, though a number of other considerations were taken into account. The S-B was included primarily as an established intelligence test to serve as a baseline in evaluating the validity of other predictor instruments. From a practical point of view, all the other predictor tests require less administration time and a less high skilled examiner than does the Binet. Also, previous research with these other instruments has suggested that all show considerable promise for the measurement of intellectual functioning among retardates. Finally, each of the other instruments possesses particular strengths which would suggest utility superior to that of the Binet, if its general predictive validity can be demonstrated to be at least as good as that of the Binet.

It might be argued that the tests designated in this study as criterion measures could prove valuable in predicting future learning ability, conceivably even better than those tests designated as predictor measures. However, in terms of this study it seemed justified to designate them as criterion tests, since their purpose is to measure what is being or has been learned. This group of tests falls into two categories: (1) measures of academic achievement and (2) tests which require learning within the test situation. Traditionally, measures of academic achievement have been used almost exclusively in validating intelligence tests designed for school age children. Since the primary purpose of the predictor tests is to predict academic learning ability, it seems logical to include such measures in the criterion battery. Two such measures, WRAT and NYAT, are included in the present battery.

It seemed appropriate also to include as criteria, measures which require the subject to demonstrate his ability to learn within the test situation. Four such tests are included in the battery. Two, the Coding Test from the Wechsler Intelligence Scale for Children (Wechsler, 1949) and the Modified Prognostic Reading Test (Woodcock, 1964) are paper and pencil tests requiring that the subject learn to associate various symbols. The second two tests, a paired associate task and a simple alternation task, are more traditional experimental learning tasks.

The selection of the battery of criterion tests, though somewhat arbitrary, was also weighted with other considerations. In terms of the achievement tests, it was considered desirable to include both an individual and a group test in the battery. Therefore, the WRAT was used, as it is the only generally available, standardized, individually administered achievement test currently on the market. The NYAT was selected primarily because it is the most readily available of several group achievement tests designed specifically for use with retarded children. The last four criterion measures were chosen primarily for their face validity: they require subjects to learn within the test situation, they are easily administered, and the test materials are readily available.

Reliability coefficients as reported in the various test manuals are reproduced in Table 1. In the interest of clarity only the best single estimate of reliability is reported in each instance. In cases where several alternative values were reported, the median of these values was taken as the best estimate. Reliability figures are considered in terms of estimates of internal consistency and stability as suggested by the Manual of Technical Recommendations (APA, 1954). In general, these data indicate adequate reliability for all tests for which data is available. More complete information concerning the reliability of each test is presented in Appendix A.

Validity coefficients for the predictor tests are reported in Table 2. Only validity information dealing with retarded populations is considered. Where several estimates of the validity of a given test were available, the range and median are reported. Estimates of validity based on comparisons among ability tests are reported separately from those based on comparisons with achievement criteria. Additional validity data, including data on normal subjects, are reported in connection with the reviews of the various tests in Appendix A.

Procedures

Basically, the procedure was to administer the predictor tests and then the criterion tests to all experimental subjects. The interval between predictor and criterion testing

Table 1

Reliability Reported in Test Manuals

Test	Internal Consistency	Stability
S-B	None reported	.85
ITPA	.99	.97
PPVT	None reported	.85
PTI	.90	.91
CPM	None reported	.90
PMAT	Not appropriate	.91
WRAT (Reading)	None reported	.95
WRAT (Spelling)	None reported	.90
WRAT (Arithmetic)	None reported	.88
NYAT (Reading)	.96	None reported
NYAT (Mathematics)	.82	None reported
NYAT (Core)	.81	None reported
Coding	Not appropriate	.60
MPRT	No manual	No manual
PAT	No manual	No manual
SAT	No manual	No manual

ranged from one to three weeks. The individually administered predictor tests were administered by a group of six trained examiners over a period of three days following standard

Table 2

Validity of Scores on Predictor Tests

with Mental Retardates

Test		Criterion Measures	
		Achievement Tests	Other Intelligence Tests
S-B			
	median	.68	.72
	range	.44-.72	.50-.78
ITPA			
	median	None reported	.70
	range		.63-.76
PPVT			
	median	.52	.66
	range	.04-.66	.27-.88
PTI		None reported	.44
CPM			
	median	None reported	.62
	range		.50-.86
PMAT			
	median	.69	.69
	range	.49-.78	.56-.82

procedures as given in the various test manuals. In the first session, the PPVT was administered, followed by the S-B; the second session included the CPM and the ITPA; and in the final session the PTI was administered. The primary purpose of such a distribution of test administrations was to minimize the likelihood of subjects becoming fatigued. While it would probably have been desirable to administer the tests in counterbalanced order, the administrative problems of such a procedure with the number of subjects and examiners involved made this impractical. Since the primary purpose of the study was to provide correlational information, order effects should not affect the results appreciably.

The order in which the tests were given is based on several considerations. The PPVT was placed first because it is generally considered to be a good rapport builder. In addition, the simplicity of format and administrative procedures suggests that PPVT scores should be less affected by initial subject resistance than those from the other instruments. Since the inclusion of subjects in the study was dependent on their obtaining S-B IQ scores between 50 and 80, economy dictated that it be administered early in the battery. Since the PPVT and S-B can be completed in one sitting, they were placed first in the testing schedule. The order of the remaining tests was determined by the toss of a coin, and the ITPA and CPM were grouped to decrease the number of class interruptions for each subject. Where possible, all tests on any one subject were administered by one examiner. The PMAT, a group test, was administered to groups of five to 15 subjects by a different examiner. Again standard procedures for administration were followed except that the testing was generally completed in two sessions on the same day.

Scheduling of criterion testing was somewhat more complicated because of differences in administrative procedures and availability of testing material. As with the predictor testing, several testing sessions were necessary. The WRAT, MPRT, and Coding were administered in one session by trained examiners following standard procedures outlined in the respective test manuals. The NYAT was given by a different examiner to each class as a group, again following standard procedures. Limitations in the amount of available testing equipment dictated that each of the learning tasks included in the criterion battery be administered separately, each by a single examiner. All tests were scored and recorded by the examiner and then re-scored by the investigator to assure accuracy.

Analyses of Data

The principal analysis used in this investigation was the correlation of all scores from predictor and criterion tests. These scores and the source of each score are listed in Table 3. All scores on the five individually administered predictor tests are derived from norms published in their respective test manuals or taken directly from the test protocols as raw scores. As no total mental age score is available for the

Table 3

Scores Submitted to Correlational Analysis

Tests	Scores	Derivation
Stanford-Binet	Mental Age	Test Protocol
	IQ	Test Manual
Illinois Test of Psycholinguistic Abilities	Language Age (total)	Test Manual
	Standard Score (total)	Test Manual
	Language Age (subtest)	Test Manual
Peabody Picture Vocabulary Test	Mental Age	Test Manual
	IQ	Test Manual
Pictorial Test of Intelligence	Mental Age (total)	Test Manual
	IQ (total)	Test Manual
	Mental Age (subtest)	Test Manual
Coloured Progressive Matrices	Raw Score	Test Protocol
Primary Mental Abilities Test	Raw Score	Test Protocol
	MA (total)	Subtest Median
	MA (subtest)	Test Manual
Wide Range Achievement Test		
Reading	Age Equivalent	Test Manual
Spelling	Age Equivalent	Test Manual
Arithmetic	Age Equivalent	Test Manual
New York Achievement Tests		
Core	Raw Score	Test Protocol
Reading	Raw Score	Test Protocol
Mathematics	Raw Score	Test Protocol
Coding	Raw Score	Test Manual
Modified Prognostic Reading Test	Raw Score	Computed
Paired Associate Task	Raw Score	Test Protocol
Simple Alternation Task	Raw Score	Test Protocol
Criterion I	Standard Score	Computed
Criterion II	Standard Score	Computed
Criterion III	Standard Score	Computed

PMAT, two scores were included as possible alternatives, the raw score and the median MA for the four subtests. Raw scores, total number of correct responses, were also used on the NYAT, the PAT, and the SAT. To obtain one set of scores which could be compared more directly with the mental ages obtained from the various predictor tests, achievement age equivalent scores were used for the WRAT. The raw score on the Coding test was the number of correct responses plus time bonuses, where appropriate, as presented in the test manual. The raw score on the MPRT is ordinarily in terms of an error score, but for purposes of this investigation all scores were subtracted from a constant (134), the total number of possible responses, in order that all scores used in the investigation would bear a positive relationship with the subject's proficiency with the task.

In addition to utilizing scores from the various criterion tests three composit criterion scores were calculated. Standard scores (Z) were than calculated for all criterion test scores to facilitate pooling these scores. Scores designated as Criterion I, Criterion II, and Criterion III were computed by averaging standard scores of the individual criterion tests. Criterion I was the mean of achievement test scores, three scores being obtained from the WRAT and three from the NYAT. Criterion II was the mean of standard scores on the four learning tasks. In order that the overall demonstrated learning score would reflect equally the achievement test and learning task data this composite score (Criterion III) was obtained by averaging C_I and C_{II}.

Compilation of the data in this manner resulted in 38 separate scores, plus duplications, bringing to 55 the total number of variables. These data were punched on IBM cards and submitted to correlational analysis using a standard product-moment program with an IBM 7072 computer.

Following the correlational analysis of these data, the relative magnitude of the correlations between predictor measures and Criterion III, and the relative magnitude of the correlations of the predictor measures and Criterion II with Criterion I were compared. This comparison was carried out using the t test for differences in correlations computed for the same population (Walker & Lev, 1958). An alpha level of .01 was used in determining significance of differences among the various validity coefficients.

Two multiple regression analyses were carried out following the procedures outlined by Cooley and Lohnes (1962) and using the IBM 7072 computer. The first of these analyses compared only the total score of each of the predictor tests taken against Criterion III, on the assumption that in a practical testing situation test users would probably use only entire tests in evaluation. A second analysis was carried out comparing PPVT MA, CPM raw score, and all subtest scores from the PMA, PTI, and ITPA with Criterion III. This second analysis was made in an effort to find a more parsimonious set of subtests which might be used to make up a battery for the prediction of learning ability.

As a final analysis, the entire matrix of independent intercorrelations was submitted to a principal axis factor analysis (Fruchter, 1954), using Guttmann's criterion to determine significant factors. The resulting factor matrix was then rotated using a biquartimin solution and a graphical rotation. This investigation into the factor structure of this group of variables was intended to evaluate the independence of various measures, particularly in terms of differentiating predictors from criteria, and determining the uniqueness of subtests. Also, loading of subtests on different factors provides some indication of the value of multiple score tests as instruments for differential diagnosis.

CHAPTER III

RESULTS

Results of analyses of test data are reported in this chapter. All results presented in connection with this investigation are reported on the basis of the scores of the 101 subjects receiving the complete test battery.

Descriptive Data

Means and standard deviations for scores on each of the variables under consideration are reported in Table 4. Age equivalent scores for the predictor tests, exclusive of subtest scores, ranged from a low of 5-5 on the ITPA to a high of 6-10 on the PMAT. The mean raw score of 13 of the Coloure Progressive Matrices is equivalent to a mental age score of 69 months. Age equivalent scores on the subtests of the ITPA PTI, and PMAT ranged from a low of 4-6 (ITPA, Auditory-Vocal Automatic) to 7-2 (PMAT, Spatial Relations and Perceptual Speed). For the tests yielding IQ scores, the range was from 65 on the Stanford-Binet to 67 on the PPVT. Overall it is apparent that the mean scores for the group are roughly equivalent on those measures where equivalence is to be expected, with the exception of the mental age score on the PMAT.

Correlational Analyses

The central concern of this investigation was with the validity coefficients obtained by correlating each of the

Table 4

Means and Standard Deviations on Experimental Variables

Variables	$\bar{X}$	SD^a
Stanford-Binet, MA	6-0	10.63
Stanford-Binet, IQ	65	7.63
Illinois Test of Psycholinguistic Abilities, LA	5-5	9.53
Illinois Test of Psycholinguistic Abilities, Z	21.7	3.95
ITPA Auditory Decoding, LA	5-9	17.80
ITPA Visual Decoding, LA	6-3	15.33
ITPA Auditory-Vocal Association, LA	5-5	14.32
ITPA Visual-Motor Association, LA	6-0	17.73
ITPA Vocal Encoding, LA	5-3	17.15
ITPA Motor Encoding, LA	5-4	18.04
ITPA Auditory-Vocal Automatic, LA	4-6	18.24
ITPA Auditory-Vocal Dequencing, LA	5-10	21.94
ITPA Visual-Motor Sequencing, LA	5-5	13.92
Peabody Picture Vocabulary Test, MA	5-7	13.82
Peabody Picture Vocabulary Test, IQ	68	12.66
Pictorial Test of Intelligence, MA	6-0	14.47
Pictorial Test of Intelligence, IQ	66	11.60
PTI Picture Vocabulary, MA	6-4	19.03
PTI Form Discrimination, MA	5-10	20.14
PTI Information & Comprehension, MA	5-8	16.18
PTI Similarities, MA	5-11	18.04
PTI Size & Number, MA	6-4	17.98
PTI Immediate Recall, MA	6-5	20.21
Coloured Progressive Matrices, RS	13	4.35
Primary Mental Abilities Test, RS	59	24.78
Primary Mental Abilities Test, MA	6-10	9.51
PMAT Verbal Meaning, MA	6-6	8.57
PMAT Spatial Relations, MA	7-2	13.59
PMAT Number Facility, MA	6-7	8.77
PMAT Perceptual Speed, MA	7-2	16.97
Wide Range Achievement Test, Reading Age	6-5	8.62
Wide Range Achievement Test, Spelling Age	6-4	6.67
Wide Range Achievement Test, Arithmetic Age	6-8	8.35
New York Achievement Test, Reading RS	22	13.32
New York Achievement Test, Mathematics RS	10	4.38
New York Achievement Test, Core RS	26	5.96
Coding, RS	29	12.32
Modified Prognostic Reading Test, RS	92	22.27
Paired Associate Task, RS	46	19.83
Simple Alternation Task, RS	67	19.65
Criterion I, Z	50	7.76
Criterion II, Z	50	6.84
Criterion III, Z	50	6.82

a. Standard deviations of age scores reported in months.

predictor tests with each of the criterion scores, including the three composite scores. These coefficients are presented in Table 5. Of primary importance is the validity demonstrated by each of the predictor tests in terms of Criterion III, the demonstrated learning score. These coefficients ranged from .42 for the PPVT to .79 for the PMAT. Multiple t tests comparing the magnitude of these coefficients indicated that the PMAT, S-B, PTI, and ITPA were essentially equal with respect to their validity, while the CPM and PPVT, though equal to each other, were significantly less valid than the other four tests. Considering only academic achievement measures (Criterion I), the picture is similar, with the PMAT, S-B, and PTI being essentially equal in their validity and the CPM and PPVT being significantly lower. However, in this analysis, the ITPA fell in somewhat of an intermediate position in that its validity was significantly lower than that of the PMAT, essentially equal to that of the S-B, PTI, and CPM, and significantly higher than that of the PPVT. Differences in the correlations of the predictor tests with Criterion II, the composite of learning task scores, were not statistically significant. Correlations of the predictor test scores with each of the individual criteria follow the same general pattern, the descending order of magnitude of correlations being in general PMAT, S-B, PTI, ITPA, CPM, and PPVT. These correlations, though generally lower than those based

Table 5

Correlations of Predictor Tests with Criterion Scores

	PMAT	SB	PTI	ITPA	CPM	PPVT
Criterion III	.79	.76	.71	.69	.50	.42
Criterion I	.79	.73	.70	.65	.48	.32
Criterion II	.67	.69	.62	.62	.46	.48
WRAT Reading	.30	.28	.27	.26	.08	.02
WRAT Spelling	.55	.52	.48	.46	.28	.12
WRAT Arithmetic	.77	.71	.75	.67	.48	.27
NYAT Reading	.66	.53	.55	.50	.43	.23
NYAT Mathematics	.73	.66	.59	.61	.45	.31
NYAT Core	.65	.67	.59	.51	.51	.50
Coding	.63	.59	.57	.47	.47	.39
MPRT	.48	.41	.42	.48	.24	.20
Paired Associates Task	.44	.53	.44	.41	.22	.39
Simple Alternation Task	.30	.36	.26	.34	.32	.35

on the pooled criterion scores, were all positive and, for the most part, significantly different than zero. Prediction of WRAT reading scores and SAT scores tends to be least adequate among the criterion tests. Examination of the raw data suggests that this may be a result of considerable skewness of the distribution of scores on these two instruments.

A second major consideration of this study was to evaluate the validity of the learning tasks as predictors of academic achievement. These correlations are presented in Table 6. The pooled score representing performance on all four learning tasks (Criterion II) correlated 0.73 with the composite scores for achievement measures (Criterion I).

Table 6

Correlations of Learning Tasks with Achievement Scores

Achievement Scores	Learning Tasks				
	C_{II}	Coding	MPRT	PAT	SAT
C_I	.73	.58	.58	.58	.26
WRAT Reading	.41	.22	.49	.32	.09
WRAT Spelling	.60	.46	.57	.46	.16
WRAT Arithmetic	.71	.60	.54	.53	.27
NYAT Reading	.52	.41	.42	.44	.15
NYAT Mathematics	.61	.49	.43	.49	.26
NYAT Core	.54	.50	.23	.46	.28

This correlation is essentially equivalent to the coefficients obtained between the better predictor tests and achievement scores. With the exception of the simple alternation learning task, the individual learning tasks tended to be similar to subtests from the predictor tests in their correlations with achievement scores. Again, it should be noted that the low correlations observed in connection with this one test may be primarily a result of the extremely skewed distribution of scores.

A final indication of the validity of the predictor tests is the degree of relationship among these tests. These correlations are presented in Table 7. Again, it is apparent that the PMAT, S-B, ITPA, and PTI, are more valid in this respect than are the CPM and the PPVT.

Table 7

Intercorrelations of Predictor Tests

	PMAT	S-B	PTI	ITPA	CPM	PPVT
PMAT		.79	.88	.66	.59	.42
S-B			.72	.78	.49	.57
PTI				.76	.68	.48
ITPA					.58	.45
CPM						.46
PPVT						

Multiple Regression Analysis

In addition to the correlational analysis, scores for all predictor tests were submitted to a multiple regression analysis using a standard program on the IBM 7072 computer. Results indicate that only two tests, the PMAT and S-B, are necessary to predict optimally the demonstrated learning score. This analysis yielded a coefficient of .136 and a Beta weight of .4959 for the PMAT raw score and a coefficient of .267 and Beta weight of .4163 for S-B MA scores. The corresponding constant for the multiple regression equation was 22.77. The resulting multiple correlation (R) was 0.841, with a standard error of estimate of 3.72 for scores using this equation. Inclusion of additional variables in the multiple regression formula neither significantly enhanced R nor reduced the standard error of estimate.

A second multiple regression analysis was carried out in which all subtest scores of the PMAT, PTI, and ITPA, and the PPVT MA and CPM raw score were compared with Criterion III. The point of diminishing returns for the addition of variables in this analysis was less clear cut than in the previous analysis. Results of this analysis for the best combination of 1, 2, 3, 4, 5, and 6 tests are presented in Table 8. Utilization of additional variables in the prediction formula led to little further improvement in R and tended to increase the standard error. Though changes in R and the standard error are consistent through the inclusion of up to six variables, the practical value of using more than three might be questioned in view of the small changes resulting from their inclusion.

Table 8

Results of Multiple Regression Analysis of Subtest Scores

Tests	Coef.[a]	Beta	R	SE	Con.[b]
1. PMAT Number Facility	.5545	.7135	.714	4.80	6.34
2. PMAT Number Facility	.3685	.4742	.799	4.14	7.61
ITPA Aud-Voc. Association	.2058	.4321			
3. PMAT Number Facility	.2828	.3639	.823	3.93	7.53
PMAT Perceptual Speed	.1010	.2514			
ITPA Aud-Voc. Association	.1765	.3707			
4. PMAT Number Facility	.2768	.3561	.833	3.85	6.67
PMAT Perceptual Speed	.0936	.2330			
ITPA Aud-Voc. Association	.1617	.3395			
ITPA Aud-Voc. Sequencing	.0422	.1357			
5. PMAT Number Facility	.2553	.3285	.844	3.75	6.23
PMAT Perceptual Speed	.0867	.2159			
ITPA Aud-Voc. Association	.1259	.2643			
ITPA Aud-Voc. Sequencing	.0483	.1554			
PTI Picture Vocabulary	.0607	.1694			
6. PMAT Number Facility	.2424	.3119	.848	3.72	6.21
PMAT Perceptual Speed	.9477	.2359			
ITPA Aud-Voc. Association	.1084	.2277			
ITPA Aud-Voc. Sequencing	.0493	.1589			
ITPA Aud-Voc. Automatic	.0357	.0955			
PTI Picture Vocabulary	.0540	.1504			

a. Coefficient for prediction formula using raw scores.
b. Constant for prediction formula.

Factor Analysis

Intercorrelations for 35 variables obtained from the overall correlation matrix were selected for the factor analysis. Certain variables (standard scores on criterion tests, IQ scores, composite criterion scores) were deleted to eliminate duplication of variables and to eliminate variables which might affect factors because of their interdependency. Variables included in the factor analysis are reported on Table 9. Subsequent tables indicate all variables by number as listed on Table 9.

A principal-axis factor analysis was carried out using a standard program for the IBM 7072 computor. Using Guttmann's criterion for significant factors, seven factors were obtained. Results of the principal-axis solution are reported in Table 10. This solution suggests a strong general factor, and six bipolar factors.

Rotation of the factor matrix was attempted in an effort to enhance interpretation. Since the matrix of intercorrelations on which the analysis was based was made up almost entirely of positive correlations the first consideration in the choice of rotational techniques was to reduce the bipolar nature of the obtained factors. Such a technique would be more consistent with the correlational data and more easily interpreted. In addition the nature of the data logically lends one to expect a general factor to appear. A biquartimin

Table 9

Variables Included in Factor Analysis

Number	Variable
1	Stanford-Binet, MA
2	Illinois Test of Psycholinguistic Abilities, LA
3	ITPA Auditory Decoding, LA
4	ITPA Visual Decoding, LA
5	ITPA Auditory-Vocal Association, LA
6	ITPA Visual-Motor Association, LA
7	ITPA Vocal Encoding, LA
8	ITPA Motor Encoding, LA
9	ITPA Auditory-Vocal Automatic, LA
10	ITPA Auditory-Vocal Sequencing, LA
11	ITPA Visual-Motor Sequencing, LA
12	Peabody Picture Vocabulary Test, MA
13	Pictorial Test of Intelligence, MA
14	PTI Picture Vocabulary, MA
15	PTI Form Discrimination, MA
16	PTI Information and Comprehension, MA
17	PTI Similarities, MA
18	PTI Size and Number, MA
19	PTI Immediate Recall, MA
20	Coloured Progressive Matrices, Raw Score
21	Primary Mental Abilities Test, Raw Score
22	PMAT Verbal Meaning, MA
23	PMAT Spatial Relations, MA
24	PMAT Number Facility, MA
25	PMAT Perceptual Speed, MA
26	Wide Range Achievement Test, Reading Age
27	Wide Range Achievement Test, Spelling Age
28	Wide Range Achievement Test, Arithmetic Age
29	New York Achievement Test, Reading Raw Score
30	New York Achievement Test, Mathematics Raw Score
31	New York Achievement Test, Core Raw Score
32	Coding, Raw Score
33	Modified Prognostic Reading Test, Raw Score
34	Paired Associate Task, Raw Score
35	Simple Alternation Task, Raw Score

Table 10

Principal-Axis Factor Analysis--Raw Factor Solution

Variables	Raw Factors							h^2
	I	II	III	IV	V	VI	VII	
1	.85	.06	-.24	.02	-.10	-.04	-.04	.80
2	.85	.11	-.22	-.32	-.04	-.20	.07	.93
3	.42	.27	.17	.03	.20	-.39	.53	.75
4	.41	.37	-.17	-.46	.19	-.01	-.13	.60
5	.78	.01	-.27	.09	.02	-.09	-.13	.71
6	.50	.05	.33	-.17	-.14	-.40	-.17	.60
7	.57	.31	-.20	-.30	-.10	-.07	-.09	.57
8	.39	.33	-.28	-.22	-.16	.53	.18	.73
9	.48	.20	-.38	.27	.22	-.27	.19	.65
10	.38	-.27	-.37	-.46	-.34	-.20	.00	.72
11	.48	-.25	.06	-.21	.34	.14	-.32	.58
12	.55	.50	-.31	.28	.02	.14	.01	.75
13	.92	.13	.18	-.15	.16	.05	.07	.95
14	.71	.30	-.05	.15	.14	.17	.04	.67
15	.69	.06	.24	-.22	.02	.16	.33	.72
16	.74	.22	.01	-.18	.13	.10	-.15	.68
17	.75	.08	.08	-.08	.15	.02	-.18	.64
18	.77	-.04	.15	-.16	.16	-.24	-.01	.73
19	.58	.11	.45	.00	.23	.07	.17	.64
20	.68	.27	.19	-.09	-.13	.21	.13	.66
21	.90	-.11	.21	.17	-.17	.04	-.07	.93
22	.71	.18	-.02	.41	-.02	-.09	-.07	.72
23	.64	.11	.32	-.03	-.29	.24	.21	.71
24	.78	-.27	.15	.02	.00	.00	.00	.70
25	.69	-.25	.28	.13	-.27	.06	-.04	.71
26	.37	-.66	-.23	-.01	.12	.18	.28	.75
27	.62	-.62	-.13	.03	.05	.11	.14	.82
28	.85	-.32	.06	-.03	.04	-.09	-.02	.84
29	.68	-.32	.04	.08	.04	.09	.09	.59
30	.76	-.17	-.05	.14	-.08	-.22	-.20	.72
31	.71	.17	-.07	.35	-.08	.05	-.08	.68
32	.67	-.06	.15	.22	-.15	-.02	.06	.55
33	.54	-.43	-.24	-.12	-.05	.05	.11	.57
34	.58	-.14	-.31	.27	.30	.06	-.24	.68
35	.37	.15	-.16	.14	-.53	-.11	.15	.52

rotation was selected as the best technique to retain the general factor and to reduce the number of negative factor loadings. The factor pattern and factor structure resulting from this rotation are reported in Table 11 and Table 12. To further clarify the meaning of the obtained factors a second rotation was carried out using the graphic technique as presented by Fruchter (1954) for obtaining orthogonal factors. Following the same logic as presented in connection with the choice of the biquartimin rotation, the axes for the graphical rotation were so placed as to maximize the number of positive loadings rather than to maximize the loadings of individual variables on any one factor. Results of this rotation are presented in Table 13. Though the rotated factor loadings resulting from these two techniques were inconsistent in many respects, the combination at the two rotated factor matrices does allow for optimal interpretation of these factors. Also it should be noted that the generally higher loadings observed in the biquartimin solution are primarily a result of the fairly high correlations (see Table 14) among these factors. With few exceptions the overall pattern of loadings is similar in the two rotations.

In summary, 35 variables were submitted to a principal-axis factor analysis. This analysis revealed seven significant factors which were rotated using graphical and biquartimin techniques of factor rotation. Results of the principal-axis solution and both rotations were reported.

Table 11

Factor Pattern--Biquartimin Rotation

Variables	Factors						
	I	II	III	IV	V	VI	VII
1	.16	.33	-.03	.12	.53	-.05	.04
2	.13	.72	.25	-.11	.17	.06	.00
3	.02	.02	.99	-.38	.01	-.11	-.11
4	-.25	.63	.11	-.19	.07	.37	.14
5	.14	.29	-.12	.02	.66	.05	-.06
6	-.20	.41	.20	.32	-.08	.05	-.45
7	-.17	.62	.00	.01	.20	.07	.10
8	.09	.13	.02	.18	.04	-.08	.76
9	.07	.11	.24	-.54	.75	-.08	-.05
10	.32	.91	-.19	-.11	-.23	-.13	-.06
11	.22	.14	-.11	.17	.10	.58	-.09
12	-.19	-.06	-.04	-.01	.85	-.11	.32
13	.14	.14	.47	.26	.15	.25	.08
14	.00	-.11	.21	.15	.57	.09	.24
15	.29	.02	.65	.28	-.24	.05	.27
16	-.05	.26	.11	.21	.29	.30	.11
17	.02	.18	.10	.26	.33	.29	-.05
18	.14	.33	.39	.07	.11	.24	-.25
19	.09	-.27	.66	.32	-.05	.22	.03
20	.00	.00	.34	.48	.03	-.04	.29
21	.23	-.08	.07	.64	.34	-.06	-.09
22	-.04	-.16	.21	.24	.79	-.12	-.10
23	-.08	-.04	-.08	.89	.06	.00	.10
24	.40	.02	.18	.37	.15	.09	-.12
25	.30	-.12	.03	.72	.08	-.14	-.10
26	.93	-.06	.10	-.14	-.06	.03	.13
27	.83	-.01	.05	.09	.08	.04	.01
28	.44	.18	.16	.24	.20	.13	-.19
29	.52	-.09	.16	.25	.18	.04	.00
30	.17	.23	-.14	.24	.51	.00	-.31
31	.00	-.14	-.08	.32	.74	-.12	.04
32	.20	-.14	.15	.43	.30	-.16	-.06
33	.62	.27	-.01	-.01	.03	.00	.05
34	.24	-.05	-.28	-.10	.84	.26	-.06
35	-.01	.20	-.06	.24	.23	-.56	.07

Table 12

Factor Structure--Biquartimin Rotation

Variables	Factors						
	I	II	III	IV	V	VI	VII
1	.44	.80	.53	.61	.62	.06	.16
2	.39	.70	.66	.57	.91	.12	.17
3	.04	.41	.70	.17	.28	-.10	-.07
4	-.08	.35	.36	.25	.59	.37	.34
5	.43	.79	.44	.51	.62	.13	.06
6	.07	.32	.48	.56	.50	.07	-.33
7	.07	.50	.42	.42	.71	.11	.28
8	.09	.34	.25	.31	.43	.04	.77
9	.20	.67	.44	.05	.33	-.07	.06
10	.45	.16	.07	.24	.71	-.12	.02
11	.40	.32	.18	.40	.31	.63	-.02
12	-.01	.76	.43	.32	.37	-.02	.44
13	.37	.73	.82	.76	.66	.37	.17
14	.19	.75	.60	.52	.42	.21	.32
15	.34	.44	.72	.62	.50	.18	.25
16	.20	.64	.56	.61	.61	.39	.25
17	.28	.65	.55	.64	.55	.38	.07
18	.38	.59	.67	.60	.61	.30	-.14
19	.18	.42	.70	.56	.24	.32	.02
20	.15	.53	.64	.69	.51	.10	.32
21	.50	.73	.61	.87	.54	.10	-.07
22	.22	.81	.54	.55	.37	-.01	-.02
23	.16	.44	.42	.83	.44	.15	.13
24	.59	.58	.53	.68	.49	.21	-.11
25	.49	.47	.43	.77	.39	.01	-.15
26	.85	.20	.07	.12	.17	.11	.00
27	.88	.41	.24	.40	.35	.16	-.08
28	.67	.64	.54	.67	.59	.24	-.15
29	.64	.52	.42	.54	.37	.16	-.04
30	.48	.69	.42	.60	.56	.08	-.22
31	.25	.78	.47	.57	.40	.00	.11
32	.38	.58	.51	.62	.37	-.04	-.06
33	.69	.35	.20	.33	.46	.08	.03
34	.45	.68	.19	.27	.27	.33	.02
35	.11	.37	.26	.34	.41	-.49	.10

Table 13

Factor Structure--Graphical Rotation

Variables	Factors I	II	III	IV	V	VI	VII	h^2
1	.53	.63	.20	.27	.07	.03	.03	.80
2	.50	.63	.48	.07	.13	-.16	-.04	.93
3	.09	.20	.30	.20	.54	-.35	.40	.75
4	.01	.49	.55	-.17	.10	.14	-.06	.60
5	.52	.58	.10	.20	.18	.11	-.08	.71
6	.30	.02	.47	.30	.31	-.10	-.30	.60
7	.15	.57	.44	.10	.00	-.02	-.13	.57
8	.03	.55	.28	.08	-.47	.12	.32	.73
9	.18	.61	-.13	.09	.43	-.10	.15	.65
10	.46	.34	.23	-.14	-.25	-.39	-.32	.72
11	.51	.08	.26	-.20	.11	.44	-.06	.58
12	.01	.73	.02	.36	.10	.21	.18	.75
13	.54	.37	.57	.25	.23	.17	.23	.95
14	.26	.53	.22	.32	.16	.27	.27	.67
15	.42	.19	.54	.22	.02	-.02	.41	.72
16	.34	.45	.49	.14	.12	.29	.04	.68
17	.45	.34	.39	.18	.22	.29	.00	.64
18	.56	.25	.45	.12	.37	.02	.00	.73
19	.32	.00	.45	.30	.30	.18	.35	.64
20	.26	.31	.49	.42	-.06	.10	.25	.66
21	.70	.23	.25	.54	.09	.15	.03	.93
22	.35	.44	.00	.54	.29	.16	.04	.72
23	.35	.13	.45	.52	-.16	.27	-.06	.71
24	.74	.14	.24	.22	.12	.09	.08	.70
25	.66	.01	.18	.49	-.04	.06	.00	.71
26	.73	.05	-.19	-.25	-.14	-.08	.30	.75
27	.88	.11	-.06	-.05	-.06	-.02	.16	.82
28	.81	.22	.25	.18	.19	.05	.02	.84
29	.70	.15	.09	.16	.05	.08	.18	.59
30	.64	.33	.10	.30	.22	.06	-.22	.72
31	.37	.48	.02	.50	.14	.20	.06	.68
32	.50	.19	.12	.48	.12	.02	.09	.55
33	.69	.24	.02	-.09	-.12	-.04	.10	.57
34	.50	.46	-.18	.00	.22	.37	.00	.68
35	.15	.37	.01	.47	-.17	-.32	-.07	.52

Table 14

Intercorrelations of Biquartimin Factors

Factor	I	II	III	IV	V	VI	VII
I	---	.28	.11	.28	.30	.10	-.12
II		---	.48	.56	.53	.05	.20
III			---	.55	.61	.08	.06
IV				---	.53	.17	.02
V					---	.10	.13
VI						---	.09
VII							---

CHAPTER IV

DISCUSSION

Chapter IV includes discussions of the results of the investigation in terms of the descriptive data obtained, the correlational analyses, and the factor analysis.

Descriptive Data

The first step in the evaluation of the tests under consideration in this study was the inspection of means, standard deviations, and ranges of the obtained scores. Since the Stanford-Binet was included as one of the selection criteria, as well as one of the tests to be evaluated, it was assured that S-B scores would fall within the range generally defining subjects as educably mentally retarded. Because of this, and the fact that the S-B is the most well established test among the predictors, these scores were used as a baseline in evaluating scores from the other tests.

The mean age equivalent score on the PTI was equal to that on the S-B (6-0). The greater range and standard deviation of scores on the PTI can be accounted for since subjects scoring at the extremes were not eliminated from the investigation as was the case with the S-B. Though the CPM does not yield an age score, it is possible to calculate age equivalents with a fair degree of accuracy from the percentile norms provided in the manual. By finding the age level at which the mean raw score (13) falls at the 50th percentile, it was possible to estimate that this score represented a mental age equivalent of 5-9. It seems unlikely that a discrepancy of the magnitude to be of practical importance.

The mean MA obtained on the PPVT (5-7) represented a discrepancy of five months below the S-B and PTI. While this discrepancy is of a magnitude which might be of practical significance, it seems best not to carry interpretations too far in this instance. The PPVT was the first test to be administered in the battery, and therefore may have been depressed to some extent by the fact that the subject was less familiar with the test situation at this point in the investigation and perhaps less comfortable with the examiner as well. This possibility seems particularly strong since examiners noted that rapport and test motivation seemed to increase, to some extent, through the testing sessions. This result is also consistent with results of previous studies (Budoff & Purseglove, 1963; Mein, 1962) which have suggested that retardates tend to score lower on the PPVT than other individual intelligence tests.

The ITPA yielded an age equivalent of 65 months, lowest of any test in the group. While the magnitude of this discrepancy would appear to be of some practical significance, it was to be expected. Similar results have been obtained in several previous studies (Mueller, 1964; Mueller & Weaver, 1964; Semmel & Mueller, 1963). In addition, it has been widely observed (McCarthy, 1964) that language ability, though highly correlated with general mental ability and roughly equivalent for the general population, tends to be among the most seriously depressed ability areas among educable mental retardates. This is especially true in groups composed largely of cultural-familial retardates or was the case with the present experimental group.

The observed mean MA (6-10) on the PMAT must also be interpreted with caution. While all other mental age scores were based on the direct raw score conversions, the norms available for the PMAT did not provide a comparable score. It appeared that the best alternative was to take the median of the MA scores which were provided for the four subtests of the PMAT as the best estimate of total mental age. However, two serious shortcomings remained. First, the lowest possible mental age for any of the four tests is 60 months, and a fairly high proportion of the experimental group fell at or near this extreme on at least one of the subtests. Coupled with the fact that much lower mental age scores can be derived from the remaining predictor measures, this suggested that the observed discrepancy in the total MA score might be at least partially a result of floor effects working to hold the PMAT estimate of MA at a higher level. Second, it was noted that in the other tests for which subscores and total scores were derived (PTI and ITPA) the total age equivalent scores were below the median MA for their respective subtests. The fact that raw scores on the PMAT consistently correlated higher with other measures than did the MA also suggested that the subtest median is not an adequate estimate of MA.

Three of the predictor tests, S-B, PTI, and PPVT, yield IQ scores, and a fourth test, the ITPA, yields a standard score allowing estimation of a comparable score. Again using the S-B (65) as the baseline, it was apparent that the PPVT (68) and PTI (66) yielded IQ scores comparable to the Binet. The slightly higher mean IQ for the PPVT was likely a result of the lower standard deviation of IQ scores on this test. Utilizing the standard scores on the ITPA, it was possible to estimate an equivalent IQ--or, more accurately, a language quotient--for these scores. The obtained estimate in this instance was 55. This discrepancy from quotients obtained from the other tests was consistent with the low age equivalent scores observed in connection with the ITPA. The PMAT manual also provides norms for the derivation of IQ scores from raw scores. However, it was immediately apparent that these scores were not appropriate for the experimental group. First, the scores provided are ratio IQ's which could not be expected to be directly comparable to the deviation IQ's obtained from the other instruments. Second, a majority of the experimental group obtained raw scores below the range covered in the conversion table, and the manual contains insufficient data for extending the range of these tables.

Scores derived from the WRAT and the NYAT fell within the range which might be expected in view of the age and ability levels of the children in the experimental groups. Although it was not possible to determine appropriate grade equivalents from the placement of the children, observation of classes suggested that most of the experimental subjects were functioning between kindergarten and early third grade level. Mean WRAT grade equivalents ranged from 1.5 in Spelling to 1.8 in Arithmetic. These scores appeared reasonable in comparison with MA expectations (6-0) for EMR children of this age level. Mean scores for the experimental group on the NYAT were: Reading, 22; Mathematics, 10; and Core, 26. These scores were consistently below the means reported for the young group in the standardization sample, but it appears that the standardization data were derived from a slightly older group. These lower means may also reflect the fact that some items in the test appear to be specifically related to life in New York City, which might also account for the experimental group's achieving at a slightly lower level. Examination of percentile norms for the tests indicated that the mean scores on the mathematics and core tests fell at the 70th percentile for children from 7-1 to 9-0. Since the ages of the experimental group extended over a year beyond this range in chronological age, these means seem to be within reasonable limits.

Insufficient previous data are available to make any statement of expectations concerning scores on the four tests requiring learning within the text situation. Examination of scores from the Coding, MPRT, and PAT revealed a fairly normal distribution of scores and gave no reasons to suggest that these tests were not fulfilling their purpose. It was fairly apparent, however, that the SAT was less adequate in differentiating among subjects than the other tests. Examination of raw data suggests that this task is in general too easy to differentiate among subjects at this age and ability level.

Correlational Analyses

Discussion of the correlational analyses conducted in the investigation are presented in terms of the correlations of predictor test with criterion tests, prediction of achievement scores by learning task scores, intercorrelations of predictor test, and multiple correlations.

Prediction of Learning Ability

The major purpose of this investigation was to compare the empirically determined validity of the six tests designated as predictors. This comparison was carried out utilizing the correlations of raw scores or age equivalent scores from the predictor tests with similar scores from the tests of demonstrated learning ability, and with standard scores derived by pooling scores from the various criterion tests. Results were essentially similar, whether comparisons were made with the overall criterion score, with the pooled score representing the various achievement tests, with each of the individual criteria, or with other predictor tests. The picture which emerged in these comparisons suggests that the PMAT, S-B, PTI, and ITPA are the most valid of the predictor tests under consideration in predicting learning ability in young mental retardates. The validity of the CPM and PPVT was significantly less than that of the other predictor tests.

The most obvious explanation for the significantly lower validity of the CPM and PPVT is that each of these tests derives its estimate of the mental ability of the subject from responses to a single type of item, while the better predictor tests provide the subject with an opportunity to respond to a variety of tasks. Further support for this explanation can be found in the examination of the correlations of the subtests from the PTI, PMAT, and ITPA with the overall criterion score. These correlations (see Appendix B) tended to fall in the same range (40's and 50's) as did the correlations involving the PPVT and CPM.

The correlations of the predictor tests with individual criterion measures also showed certain consistencies which are worth noting. Correlations of the predictor tests with seven of the criteria (WRAT Reading, WRAT Spelling, WRAT Arithmetic, NYAT Reading, NYAT Mathematics, MPRT, and PAT) showed essentially the same pattern as their correlations with Criterion III, indicating better validity for the PMAT, S-B, PTI, and ITPA than for the CPM and PPVT. Validity of the predictor tests for the NYAT Core and Coding tests was also comparable except for the somewhat lower validity of the ITPA. Correlations with the SAT were essentially equal for all predictor tests. Overall, the predictor measures appeared to be least valid in predicting WRAT Reading and SAT scores.

Examination of the distribution of scores for these two tests suggested a possible reason for the lower coefficients on these two tests. It appears that the SAT is not sufficiently difficult to differentiate among these subjects. A total of 65 subjects obtained scores of 70 or above from a possible score of 80 on this test. It appeared that a number of children came to the testing situation with a set for alternation. This was evidenced by the fact that 12 children achieved perfect scores on this test. It was also apparent that possessing this learning set did not correlate highly with proficiency on the other learning tasks.

The low correlations found between the various predictor measures and the WRAT reading test may also have resulted in part from a somewhat skewed distribution, although this was less marked than in the case of the SAT. Examination of the distribution of the scores on the WRAT reading indicated that a fairly large proportion of the subjects were clustered near the bottom end of the scale. It may also be, however, that reading achievement among children of the age and ability level under consideration here is simply less highly correlated with general mental ability than are other areas of achievement. The fact that the lowest correlations obtained for the NYAT were also in the area of reading lends some support to this idea. Although this finding is at odds with the majority of previous research in this area, it was consistent with results obtained by Moss (1962), who also worked with a group of young, educable mental retardates. This may merely reflect the fact that these two studies dealt with children who were generally at the very beginning stages of reading, and whose reading achievement, therefore, was difficult to measure reliably. It may, however, indicate that important qualitative differences do exist between patterns of mental abilities at various mental age levels with more verbal skills being a less important aspect of academic and mental development at the lower age levels. This observation is consistent with the findings of Reed (1958) that prediction of reading at the grade one level requires different components than at higher levels.

Prediction of Achievement by Learning Tasks

Another consideration of this investigation was the possibility of utilizing tests requiring learning within the test situation as a means of predicting achievement. To evaluate this, the correlations of the learning tasks and their pooled score (Criterion II) with the achievement tests and their pooled score (Criterion I) were considered. Although correlations involving the individual learning tasks were not particularly high, Criterion II, a score computed by pooling scores from the four learning tasks, held up very well as a predictor of achievement test scores. The obtained correlation of 0.73 with the pooled achievement score indicated validity equal to that of the Binet and not significantly different from that of the PMAT, the best single predictor of this criterion score. Furthermore, Criterion II proved to be the best predictor of WRAT Reading and Spelling scores, and coefficients obtained against the WRAT Arithmetic and the NYAT were in the same range as those of the better predictor tests. With the exception of the SAT, the individual tasks making up Criterion II also appeared to show promise as predictors of achievement. Correlations between SAT scores and the various scores, though positive, were generally low. The reason for this appeared to be lack of differentiation of subjects on this test, as discussed in connection with its correlations with the predictor measures. Correlations of the Coding, MPRT, and PAT with the pooled achievement score were equal, 0.58. Overall these three tasks compared favorably with other single channel tasks in the prediction of achievement scores. The Coding test appeared to be somewhat superior to the CPM and PPVT in the prediction of achievement test scores, but less adequate than the more complex predictor measures. The pattern of correlations involving the PAT was essentially the same as that of the Coding test. The correlation of the MPRT with the pooled score for the achievement measures was identical with that of the Coding and PAT. However, the pattern of correlations with individual measures of achievement was somewhat different. This test proved to be the best single predictor of WRAT Reading and WRAT Spelling scores, being significantly superior to all the predictor measures in its correlation with WRAT Reading. However, the correlation of the MPRT with the remaining achievement tests was more in line with the poorer predictor tests. Overall, the results of this investigation suggested that tasks such as these, requiring the subject to demonstrate his ability to learn within the test situation, show considerable promise as predictors of academic achievement.

Intercorrelations Among Predictor Tests

Evaluation of the intercorrelations among the predictor tests provided further evidence for their validity. Again it seemed appropriate to utilize the S-B as a baseline. Correlations with the S-B were 0.78 for the ITPA, 0.72 for the PTI, and 0.70 for the PMAT. In comparison with previous studies of test validity (see Appendix A), these coefficients may be taken as indicating adequate validity. It should also be noted that the magnitude of the differences in these coefficients was not statistically significant. Correlations with the S-B of 0.57 for the PPVT and 0.49 for the CPM indicated that these instruments are significantly less valid as predictors of S-B scores. Examination of the correlations among all the predictor tests led to the same conclusion. Correlations among the S-B, PTI, ITPA, and PMAT ranged from 0.66 to 0.78 while correlations involving the CPM and PPVT ranged from 0.42 to 0.68. These results indicate that the CPM and PPVT are consistently less valid than the other instruments in terms of comparisons among intelligence tests.

Multiple Prediction

Two multiple regression analyses were carried out in connection with this investigation. In the first, total scores from tests designated as predictors were utilized, and in the second, only scores from the two very short predictor tests (CPM and PPVT) and subtest scores from the ITPA, PTI, and PMA. In both analyses, Criterion III was the predicted score.

Results of the first multiple regression analysis were very clear-cut. A multiple correlation of 0.84 was obtained utilizing only two test scores, the raw scores from the PMAT, and the MA score from the S-B. This analysis suggested that the practice currently used in our public schools should be adequate for determining placement of young retarded children. This practice, which is in fairly general use, is to administer a group intelligence test as a first screening device and to follow this up with the administration of an individual test. The results of this investigation indicated that this procedure, utilizing the PMAT as the group test and the S-B as the individual test, should be an adequate placement criterion. On the basis of this study, it was not possible to state that the PMAT was superior to other group tests for this purpose, since no other group tests of intelligence were included in the present study. However, previous research in the area has rarely shown any group test to compare as favorably with individual intelligence tests as did the PMAT in this investigation. These results do suggest that the S-B is the best instrument among those individual tests included in this investigation for use in this situation, although the substitution of the PTI or ITPA would not seriously reduce the level of prediction.

Although it seems somewhat unlikely that in general practice psychological examinations would be carried out utilizing parts of a number of different tests rather than complete tests, the possibility of doing so has certain merits. Certainly it is of theoretical interest to examine the number and types of activities which appear to be required to predict demonstrated learning ability optimally. Therefore, a second multiple regression analysis was carried out, utilizing all of the obtained predictor scores which were derived from tests involving a single type of item and the overall criterion score. Predictor scores in this analysis included the raw score from the CPM, the MA from the PPVT, MA scores from each of the subtests of the PMAT and the PTI, and LA scores from each of the subtests of the ITPA. In terms of these variables, it was fairly clear that at least three subtests must be administered to obtain optimal prediction of a demonstrated learning score. Addition of fourth, fifth, and sixth subtests continued to improve prediction, but whether this slight improvement was significant from a practical standpoint was questionable. Inclusion of additional variables did not enhance prediction. The three tests yielding the best multiple prediction of the overall criterion were Number Facility and Perceptual Speed from the PMAT, and Auditory-Vocal Association from the ITPA. This suggested that these tests measured the three most important aspects of learning ability among children at this mental age level. This seems fairly logical when one analyzes the content of the three tests. On the basis of the content of these tests it appears that verbal ability, number ability, and perceptual ability are the primary components of learning ability for young EMR children. This combination is consistent with the curriculum of many special classes at this level. Auditory-Vocal Association consists of items requiring the subject to complete an opposite analogy such as: Brother is a boy, sister is a ________. Previous studies of the ITPA have indicated that this subtest tends to be the best single measure of overall verbal ability and also to be the most highly correlated of the ITPA subtests in comparisons with mental age. The fact that this test showed up among the best predictors of learning ability was consistent with the emphasis placed on language development in classes for retarded children and with the practice of including measures of verbal intelligence in most test batteries for prediction of learning ability. Number facility and perception, which are measured by the PMAT subtests, are also of obvious importance in the curriculum of the primary EMR class.

While the statistical analysis indicated only slight improvement in the multiple prediction of learning ability through the inclusion of additional tests, examination of the tests which would be included suggested that these tests may be of some value. The fourth test which would be added would be Auditory-Vocal Sequencing from the ITPA. This test is a modification of the digit span technique found in many intelligence tests. Assuming that this test primarily measures short term retention (Baumiester & Bartlett, 1963) it seems reasonable to include this test, since it appears to be concerned with a relatively independent facet of overall intellectual functioning. The fifth test to be included in a composite battery from those subtests under consideration would be the Picture Vocabulary Test from the PTI. Vocabulary tests have had fairly wide acceptance among authors of intelligence tests as measures of general mental ability, and thus the inclusion of this subtest seems reasonable. However, the extent to which it is independent of the abilities tapped by the Auditory-Vocal Association subtest was probably reflected in the very slight gain in multiple prediction, which resulted from the inclusion of this test. The sixth test, Auditory-Vocal Automatic from the ITPA, is essentially a test of knowledge of grammatical structure in language. This seems at least somewhat analogous to subtests in other batteries which are intended to measure the extent of a person's contact with and knowledge of his environment. However, this is a highly verbal task, and its independence from the Auditory-Vocal Association subtest was probably reflected again in the very small increase it contributed to multiple prediction.

On the basis of these results, the recommendation of this author would be to include the best combination of four subtests to make up a battery designed to predict learning ability in young EMR children. As indicated, these tests would be the Number Facility and Perceptual Speed subtests from the PMAT, and the Auditory-Vocal Association and Auditory-Vocal Sequencing tests from the ITPA. The value of this suggestion of course would need to be cross-validated through the administration of tests to another sample similar to this experimental group. Another interesting follow-up of these observations would be to administer an alternate group of tests analogous to this set in an effort to determine the extent to which the rather high level of prediction observed in this study was a function of the individual tests used, or more generally, the areas of intellectual functioning being tapped.

Factor Analysis

Interpretation of the results of the factor analysis are based primarily on the factor structure revealed by the two rotations of the principal-axis factor matrix. However, one inference from the raw solution seems warranted. This solution gave the best estimate of the strength of the general factor for those variables under consideration. This solution revealed an extremely strong general factor with loadings for the 35 variables ranging from 0.37 to 0.92. Although these loadings were all positive and fairly high throughout, the highest loadings on the factor were obtained for variables representing overall estimates of mental development including the Stanford-Binet, ITPA, PTI, and PMAT. Single channel tests which also obtained high loadings on this general factor included Auditory-Vocal Association from the ITPA, the Coloured Progressive Matrices, WRAT Spelling and Arithmetic, Coding, and essentially all the subtests from the PTI, PMAT, and the NYAT.

Naming and interpretation of factors from the rotated matrices was relatively difficult because of the large number of variables showing high loadings on each of the factors and because of the lack of proven reference variables. The differences which existed between the two rotations also led to some difficulty in interpretation, but it seemed preferable to consider these two rotations as different estimates of the same factors since both were based on the same principal-axis factors.

Factor I was fairly clear-cut with essentially the same variables loading heavily in both rotations although the overall level of the loading was, of course, considerably higher in the biquartimin rotation. This factor was interpreted as an academic achievement factor and named General Academic Achievement on the basis of consistently high loadings on five of the six measures of achievement included in the analysis. Other variables loading very highly on this factor were the PMAT total score, Number Facility, and Perceptual

Speed. This would suggest that number facility, and perceptual speed tend to be among the most important aspects of academic achievement among children at this age and ability level, a finding consistent with the work of Reed (1958). Although the loadings of other variables included in the analysis were generally lower, the majority of these variables still maintained moderate loadings on this factor, which was consistent with the fact that these measures were designed as predictors of academic achievement.

Factor II was also interpreted as a general factor which was called General Verbal Ability. This interpretation was based on the particularly strong loadings of the S-B and ITPA on this factor. Other tests with high loadings on Factor II included the association, encoding, and automatic subtests from the ITPA, the PPVT, and PMAT Verbal Meaning. The content of these tests was consistent with interpretation of the factor as General Verbal Ability. The high loadings of the PAT, CPM, and Visual Decoding from the ITPA on this factor was likely a result of the need for verbal mediation in responding to these tasks. The high loadings of Auditory-Vocal Automatic from the ITPA, the New York Core, and the three picture vocabulary tests suggested that this factor might also be related to aculturation, as these tasks seemed to be particularly high in content of a cultural nature. This observation is consistent with the fact that tests which have been designed to be culture free or culture fair (Leiter, Nebraska, Raven's) have tended to avoid highly verbal test items.

Factor III could also be interpreted as a fairly general intellectual factor, but more heavily loaded with perceptual aspects of intelligence as opposed to the more verbal and cultural aspects associated with Factor II. Thus, this factor might be called General Perceptual Ability. This interpretation was based on the high loadings of the CPM, the PTI and its subtests, the PMAT and its Spatial Relations subtest, and the Coding test. Although Factor III was interpreted as General Perceptual Ability, it was apparent from the pattern of factor loadings that measures involving visual perception were more closely related to the factor than those requiring auditory perception. This may indicate that vision is the more important channel of perception at this level of age and ability.

Factor IV was interpreted as Association. Again visual perception appeared to be involved, but the pattern of loadings here suggested that the critical ability being tapped was the association of symbols rather than perceptual ability. High loadings on the ITPA association subtests, all of the PMAT and PTI subtests, the New York Core, and Coding support this conclusion.

Measures with high loadings on Factor V included Auditory

Decoding, Auditory-Vocal Association, and Auditory-Vocal Automatic from the ITPA; the three picture vocabulary tests; the PAT; and Information and Comprehension, Similarities, and Size and Number from the PTI. With the exception of the PAT these tests all require a high level of attention to auditory stimuli. The PAT and Immediate Recall from the PTI, and Coding require careful attention to visual stimuli. On the basis of these measures this factor was interpreted as Attention, although whether it represents a general attention factor or a more specific auditory attention factor is not entirely clear since several tests (Visual-Motor Sequencing, MPRT, Spatial Relations, and Perceptual Speed) which require close attention to visual stimuli have relatively low loadings on the factor.

Visual-Motor Sequencing was by far the highest loading in both rotations of Factor VI, indicating that this factor may be involved with sequential learning effects. However, two loadings (PAT and Immediate Recall) suggested that this factor may also be closely related to short term memory. Since the sequencing tests also require memory, this factor was interpreted as Short Term Memory.

Factor VII was defined by Motor Encoding in the biquartimin rotation and on the graphical rotation by a number of measures which apparently had little in common (Auditory Decoding, Form Discrimination, Immediate Recall, Motor Encoding, and WRAT reading). In view of the appearance of Motor Encoding in both rotations, this factor might be interpreted as Motor Encoding, but in view of the loadings of other measures it seemed best to leave this factor uninterpreted.

As might be expected for a test of very general content, the S-B loaded on several factors. Results of the graphical rotation indicated high loadings on General Academic Achievement and General Verbal Ability. The biquartimin solution revealed high loadings on all the interpreted factors with the exception of Memory. Factor loadings of the ITPA were similar to those of the S-B although its loadings indicated its more verbal nature. The PPVT showed a strong loading only on General Verbal Ability, although it loaded moderately on Association in both rotations and on Attention and General Perceptual Ability on the biquartimin rotation. The pattern of loadings for the PTI and PMAT also revealed a close relationship to all the interpreted factors, reflecting their general nature. The PMAT achieved the highest loadings on General Academic Achievement. In general, these results were consistent with what might be expected from logical analysis of all tests.

The achievement tests tended to load together and defined the factor named General Academic Achievement. The MPRT also showed its highest loading on this factor. The New York Core Achievement Test, however, did not load with the other achievement tests, but was most closely related to General Verbal Ability and Association. Coding loaded primarily on Association with moderate loadings on the verbal and perceptual factors. Again these results were consistent with logical analysis of the test content. The primary loading of the PAT was on General Verbal Ability with a moderate loading on Short Term Memory. The SAT loaded strongly only on Association on the graphical rotation.

Examination of the loadings of the subtests from the ITPA, PTI, and PMAT revealed little independence among subtests except for the tests of the ITPA. The tests of the ITPA showed a fair degree of independence, although all loaded at least moderately on General Verbal Ability. All but Motor Encoding and Visual-Motor Sequencing loaded on Attention. Loadings of the subtests of both the PTI and PMAT were too general to indicate significant independence except in the case of the Spatial Relations test of the PMAT. This test loaded specifically on Attention and General Perceptual Ability.

Overall, results of the factor analysis and subsequent rotations were consistent with logical analysis of the abilities tapped by the various tests; however, some incongruities were noted. Perhaps the most notable was the general nature of the subtests of the PMAT. Results of this investigation strongly suggested that these tests measure general intelligence instead of the independent factors they were intended to measure. This need not be viewed as a criticism of the PMAT, however, since results reported previously indicated that the test predicts general achievement very well. The subtests of the PTI might also have been expected to show a higher degree of independence. The failure of the New York Core to show high loadings on General Academic Achievement might also be viewed as an incongruity. The pattern of loadings for this test corresponded better to those of the intelligence tests than to other achievement measures. The high loading of the MPRT suggested that this test is not as independent from measures of academic achievement as was hoped for tasks included as measures requiring learning within the test situation. Finally, the virtual lack of high loadings of the SAT may be interpreted as further evidence of this task's inadequacy for measurement of young educable retardates.

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

This investigation was designed to evaluate the empirical validity of six general ability tests, designated predictors, against a number of measures of learning. The six predictor tests were the Stanford-Binet, Illinois Test of Psycholinguistic Abilities, Peabody Picture Vocabulary Test, Pictorial Test of Intelligence, Coloured Progressive Matrices, and Primary Mental Abilities Test. Demonstrated learning ability was defined as performance on two achievement tests and four tasks requiring learning within the test situation. These criteria were the Wide Range Achievement Test, the New York Achievement Tests, the Coding subtest from the WISC, a modification of the Prognostic Reading Test, a paired associate learning task, and a simple alternation learning task. The study was also viewed as an investigation of the use of multiple criteria in the validation of ability tests.

The investigation was carried out utilizing a sample of 101 educable mentally retarded children between the ages of seven and 11 years. All tests were administered by qualified examiners over a period of approximately three weeks. Analyses of the collected data included correlation of all variables, multiple regression analysis of predictor test scores on the overall criterion score, and factor analysis of selected variables.

Validity coefficients of the predictor tests against the overall criterion score were PMAT, 0.79; S-B, 0.76; PTI, 0.71; ITPA, 0.69; CPM, 0.50; and PPVT, 0.42. Comparison of these coefficients indicated that the PMAT, S-B, PTI, and ITPA were significantly more valid than the CPM and PPVT. Further comparisons of correlations of the predictor tests wich individual criterion tests yielded similar results. It was also observed that a composite score developed by pooling scores of the four learning tasks was essentially equal to the better predictor tests taking achievement test scores as the criterion.

Multiple regression analysis indicated that optimal prediction of the overall criterion score required administration of only two predictors, the PMAT and S-B. A second multiple regression analysis suggested that three to six short, single channel tests were required for optimal prediction of the overall criterion. These tests in descending order of importance were PMAT Number Facility, ITPA Auditory-Vocal Association, PMAT Perceptual Speed, ITPA Auditory-Vocal Sequencing, PTI Picture Vocabulary, and ITPA Auditory-Vocal Automatic.

Factor analysis of 35 selected variables was carried out using a principal-axis solution and both a biquartimin and graphical rotations. Seven significant factors were obtained in this analysis. One factor was uninterpreted and the remaining six were identified as General Academic Achievement, General Verbal Ability, General Perceptual Ability, Association, Attention, and Short Term Memory. With the exception of the New York Core all achievement tests loaded primarily on General Academic Achievement as did the MPRT. The S-B, ITPA, PTI, and PMAT all loaded at least moderately on several factors but tended to correlate highest with the more general factors. The subtests of the PTI and PMAT also showed very general loadings while those of the ITPA exhibited somewhat greater independence.

On the basis of this investigation it was concluded that the PMAT, S-B, PTI, and ITPA were valid predictors of learning ability among primary age EMR children, while the CPM and PPVT were significantly less valid. Results of the multiple regression and factor analyses indicates that the most important factors in the intellectual functioning of this group were verbal, numerical, and perceptual ability.

Several implications for psychometric practice are suggested by the results of this investigation. Perhaps most important is the support for the practice of using a combination of one group test and one individual test of intelligence in evaluation of children suspected of being mentally retarded. Also, although the multiple regression analysis indicates that the S-B and PMAT allow the highest level of prediction among the tests studied, the correlational analysis suggests that prediction based on the use of the PTI or ITPA is not significantly lower. Therefore, where a more diagnostic test of mental ability is desirable, either of these tests could be substituted for the S-B without an appreciable loss in prediction. This investigation also suggests that neither the CPM nor the PPVT is an adequate substitute for the more comprehensive instruments for the purpose of predicting general learning ability for educable mental retardates in the early school years. This is not meant to indicate that these tests could not be of value for other purposes or for other populations.

Results of this investigation also suggest implications for future research in test development and validation. The techniques of using multiple criteria and direct comparisons among validity coefficients were explored and appear to be adequate. Two possibilities for further test development were suggested by data from this study. First, the battery of learning tasks proved quite valid for the prediction of achievement. Second, the data suggest that prediction of learning ability at this level of age and ability depends primarily on measurement of verbal, numerical and perceptual ability. Both these findings suggest possible bases for development of new instruments or modification of existing tests. In view of the high validity of raw scores from the PMAT it seems imperative that more adequately derived scores be developed. Finally, additional research is needed to verify the results of this investigation and to extend the study over a broader range of age, intelligence, and school achievement.

REFERENCES

References

American Psychological Association, American Educational Research Association, and National Council on Measurement Used in Education, Joint Committee of. *Technical recommendations for psychological tests and diagnostic techniques*. Washington, D. C.: Amer. Psychol. Assn., 1954.

Anastasi, Anne. *Psychological testing*. (2nd ed.) New York: Macmillan, 1961.

Binet, A., & Simon, T. Methodes nouvolles pour le diagnostic du niveau intellectue des anormaux. *Annee psychol*., 1905, 11, 191-244.

Budoff, M., & Purseglove, E. M. PPVT performance of institutionalized mentally retarded adolescents. *Amer. J. ment. Defic*., 1963, 67, 756-760.

Cooley, W. W., & Lohnes, P. R. *Multivariate procedures for the behavioral sciences*. New York: Wiley, 1962.

Denny, M. R. Learning and performance. In H. A. Stevens & R. Heber (Eds.), *Mental retardation*. Chicago: Univer. of Chicago Press, 1964. Pp. 100-142.

Dunn, L. M. *Peabody picture vocabulary test*. Minneapolis: American Guidance Service, 1959.

French, J. L. *Pictorial test of intelligence*. Boston: Houghton Mifflin, 1963.

Fruchter, B. *Introduction to factor analysis*. Princeton: Van Nostrand, 1954.

Jastak, J. Wide range achievement test. New York: Psychological Corp., 1946.

Kirk, S. A. Research in education. In H. S. Stevens & R. Heber (Eds.), Mental retardation. Chicago: Univer. of Chicago Press, 1964. Pp. 57-99.

McCarthy, J. J., & Kirk, S. A. The Illinois test of psycholinguistic abilities. Madison, Wisc.: ITPA, Inc., 1961.

McCarthy, J. J., & Olson, J. L. Validity studies on the Illinois test of psycholinguistic abilities. Madison, Wisc.: Authors, 1964.

Mein, R. Use of the Peabody picture vocabulary test with severely subnormal patients. Amer. J. ment. Defic., 1960, 65, 482-485.

Moss, J. W. An evaluation of the Peabody picture vocabulary test with the PMA and 1937 Stanford-Binet. Unpublished paper, Univer. of Illinois, Institute for Research on Exceptional Children, 1962.

Mueller, M. W. Comparison of psycholinguistic patterns of gifted and retarded children. J. school Psychol., in press.

Mueller, M. W., & Weaver, S. J. Psycholinguistic abilities of institutionalized and non-institutionalized trainable mental retardates. Amer. J. ment. Defic., 1964, 68, 775-783.

Raven, J. C. Progressive matrices (1947). London: H. K. Lewis and George G. Harrap, 1947.

Reed, J. C. The relationship between the primary mental abilities and reading achievement at given developmental levels. *Amer. Psychologist*, 1958, 13, 324.

Semmel, M. I., & Mueller, M. W. A factor analysis of the Illinois test of psycholinguistic abilities with mentally retarded children. Unpublished paper, Peabody College, 1963.

Terman, L. M., & Merrill, Maud A. *Stanford-Binet intelligence scale: manual for the third revision, Form L-M.* Boston: Houghton Mifflin, 1960.

Thurstone, L. L., & Thurstone, Thelma G. *Examiner manual, PMA, primary mental abilities for grades 2-4.* Chicago: Science Research Associates, 1962.

Walker, Helen M., & Lev, J. *Elementary statistical methods.* (Rev. Ed.) New York: Holt, Rinehart, Winston, 1958.

Wechsler, D. *Wechsler intelligence scale for children.* New York: Psychol. Corp., 1949.

Woodcock, R. W. Modified prognostic reading test. Unpublished material, Peabody College, 1964.

Wrightstone, F. W., Forlano, G., Lepkowaski, J. R., Sontag, M., & Edelstein, J. D. A comparison of educational outcomes under singletrack and two-track plans for educable mentally retarded children. U. S. Office of Education, Cooperative Research Branch, Contract No. 6908, Project No. 144, 1959.

APPENDIXES

APPENDIX A

INSTRUMENTS USED IN THE STUDY

APPENDIX A

INSTRUMENTS USED IN THE STUDY

Appendix A includes a brief review of each of the instruments used in this investigation. Each test is reviewed in terms of description, standardization, reliability, validity, use with mental retardates, strengths and weaknesses, and rationale for inclusion in the study.

Stanford-Binet Intelligence Scale

The S-B (Terman & Merrill, 1960) is a standardized, individual intelligence test yielding scores for mental age and deviation IQ's. Items are grouped into age levels, according to difficulty, ranging from age two to superior adult, and ranging from simple manipulation of objects to abstract reasoning. Although a number of performance type items are included, particularly at lower age levels, the test is essentially verbal in nature. Reliability and validity of earlier editions of the scale, particularly the 1937 revision, have been extensively studied. Alternate form reliabilities for the 1937 revision of the S-B range from 0.83 to 0.98 depending on age and IQ level, higher correlations being obtained at upper age levels and at low and average IQ levels. Test-retest reliabilities reported by the author are also quite adequate. Studies using relatively short (one year or less) interest intervals have generally obtained reliability coefficients in the .80's and .90's (i.e., Sontag _et al_., 1958). Bradway, _et al_. (1958) reported a reliability of 0.65 following a 10-year interval and 0.59 with an interval of 25 years, for scores of children in the standardization sample who were from 2.0 to 5.5 at the time of original testing. These studies are fairly representative of reliability studies of the 1937 S-B. Validity in predicting school achievement has been generally good, particularly in predicting achievement in more verbally oriented subjects such as language and reading. A typical study (Bond, 1940) reported correlations from 0.43 to 0.73 between Binet scores and achievement in various school subjects among tenth grade children. Although a restandardization of the scale was not carried out in connection with the 1960 edition, the authors suggest that this revision retains the main characteristics of the 1937 revision including high reliability and validity.

Studies of the S-B with retardates have been primarily with earlier editions of the test. Collman and Newlyn (1958) reported test-retest reliability of 0.93 on a group of 182 regardates over an inter-test interval of one year. Stott (1960) reported a test-retest reliability of 0.80 for S-B scores of a group of 942 retardates. Inter-test interval was not held constant in this study, ranging from one month to 10 years. Silverstein (1963) reported a reliability of 0.95 for the 1960 S-B, but his source for this figure is not clear.

Miller (1960) in a review article suggests that research evidence concerning the use of the S-B with retardates indicates a high degree of stability of IQ scores over time, particularly when standard scores are compared. One validity study (Rohns & Haworth, 1962) on the 1960 S-B reported correlations of 0.69 with WISC verbal IQ and 0.72 with WISC full scale IQ. Moss (1962) reported correlations for the S-B with several measures of achievement ranging from 0.44 to 0.68. Validity of the S-B has also been evaluated in terms of correlations with Vineland social age (SA). Goulet and Barclay (1963) reported on correlations of 0.78 between MA and SA.

On the basis of the broad acceptance of earlier editions of the S-B, and the need for data specifically related to the 1960 edition, it seemed vital to include the S-B among the tests to be evaluated in this study. Although various authors (i.e., Anastasi, 1961) critized the Binet on specific points (highly verbal nature of the test, unsuitability for measurement of differential aptitudes, weakness with adult subjects), it is likely that fairly general agreement could be obtained that it is the best single instrument for evaluating the verbal intelligence of children at the age and ability level being considered in this study. Certainly the reports of Silverstein (1963b) and Weise (1960) indicate that it is among the most widely used.

Illinois Test of Psycholinguistic Abilities

The ITPA (McCarthy & Kirk, 1961) is an individually administered test measuring language functioning, in terms of the Osgood (1957) language model, among children from 2-6 to 9-0 years of age. It yields age equivalent and standard scores on total language functioning and on each of nine subtests: Auditory Decoding, Visual Decoding, Auditory-Vocal Association, Visual-Motor Association, Vocal Encoding, Motor Encoding, Auditory-Vocal Automatic, Auditory-Vocal Sequencing, and Visual-Motor Sequencing. It was standardized on a total group of 700 children of average intelligence.

Reliability is exceptionally high, a split-half reliability of 0.99 and a test-retest reliability of 0.97 being reported for the standardization sample. At present, evidence of validity of the Illinois Test of Psycholinguistic Abilities is scant. Early studies of the test have indicated fairly high correlations with tests of general intelligence. In the standardization of the test (McCarthy & Kirk, 1961) a correlation of 0.96 was obtained between age scores derived from the Binet and the ITPA. Sievers (1955) obtained a correlation of 0.66 between MA and total scores on an earlier version of the ITPA for a group of normal children. McCarthy and Olson (1964) reported an extensive study of the validity of the ITPA with a group of 86 children ranging in age from 7-0 to 8-6. They concluded that the concurrent, construct, and predictive validities of the ITPA are adequate. Content and diagnostic validities were reported as less adequate, but this appeared to be as much a function of the author's obtaining less information in this regard as of actually demonstrating lower validity.

A number of studies (Kirk & Bateman, 1962; Kirk, Cass & Bateman, 1962) have found the test useful in planning individual programs of instruction for the remediation of learning disabilities. While these studies did not result in obtaining validity coefficients, they do give some support to the validity of the test.

Several studies dealing with retarded subjects have been reported. Mueller and Smith (1963) reported a test-retest reliability of 0.94 on tests administered to 13 educable retardates with an inter-test interval of nine months. Semmel and Mueller (1963) obtained a correlation of 0.63 between age scores derived from the ITPA and the S-B, and Mueller and Weaver (1963) a correlation of 0.76. Both these correlations were obtained from tests given to mentally retarded subjects and may have been depressed somewhat by the restricted range of the sample.

A comprehensive analysis of the strengths and weaknesses of the ITPA must await further research. On the basis of initial investigations the test appears to be highly reliable, to have at least adequate validity, and to be useful in differential diagnosis of learning disabilities. Weaknesses include its rather narrow range and insufficient validity data.

The ITPA is included in the present study primarily on the basis of the promise it has shown in early studies and the extensive research its publication has stimulated. In addition to observations from studies of the ITPA, the experience of the present author in working with the test has suggested that the ITPA may be particularly valuable in predicting the learning ability of retardates. This is based on the fairly consistent correlations between ITPA scores and measures of intelligence, the fact that descriptions of children based on their ITPA performance have been accepted by teachers as highly accurate, and that the language age scores obtained from the ITPA have seemed to reflect the academic performance of mentally retarded children to a greater extent than have MA scores for the same children.

Peabody Picture Vocabulary Test

The PPVT (Dunn, 1959) is an individually administered test yielding an estimate of intelligence based upon the measurement of hearing vocabulary. The subject is required to indicate which of four response pictures correctly depicts a stimulus word presented orally by the examiner. There are 150 items on the total test, but it is necessary to administer only a critical span from the point where a subject makes eight consecutive correct responses to the point where correct responses are down to a chance level (six of eight consecutive responses incorrect). The PPVT is appropriate for persons from age two years-three months to age 18 years, is available in two forms, and generally requires only 10 to 15 minutes to administer. Derived scores can be obtained for MA, IQ, and percentile rank.

The test was standardized on approximately 4,000 Caucasian subjects in the Nashville, Tennessee area. The test manual (Dunn, 1965) reports alternate form reliability coefficients ranging from 0.67 to 0.84 for various age levels across the full spectrum of intelligence with a median value of 0.77. The manual also reports validity coefficients ranging from 0.71 to 0.94 in comparisons with various other intelligence tests and coefficients ranging from 0.52 to 0.90 between PPVT scores and teacher ratings of achievement for samples of mentally retarded and cerebral palsied children.

Subsequent work with the PPVT has led to numerous questions regarding its validity, and to some extent, its reliability. In studies of the validity of the PPVT dealing with normal children across a broad range of intellect (Klaus & Starke, 1964; Lavitt, 1963; Lindstrom, 1961; Tempero & Ivanoff, 1960), correlations with various other measures of intelligence have been reported ranging from 0.36 (Lindstrom, 1961) to 0.82 (Lavitt, 1963), with the majority of the obtained coefficients falling between 0.60 and 0.80. While many of these coefficients may be regarded as low, they represent significant correlations in all cases. Lindstrom (1960) also observed that IQ scores derived from the PPVT averaged 10 points above those derived from the WISC.

Several studies dealing specifically with estimates of the reliability of the PPVT with retarded groups have been reported. Budoff and Purseglove (1963) reported a correlation of 0.85 for equivalent forms and 0.87 for stability of MA scores over an inter-test interval of one month. An alternate form reliability of 0.86 was reported by Kimbrell (1960) for PPVT IQ scores. Moss (1962) obtained coefficients of 0.68 for IQ scores with a time lapse of three years and 0.54 over an interval of two years. A large number of studies have dealt directly with the question of the validity of the PPVT with mentally retarded subjects. Budoff and Purseglove (1963) found correlations ranging from 0.83 to 0.88 between the two forms of the PPVT and two forms of the S-B administered to a group of 46 institutionalized, adolescent retardates. In this study the mean MA on the PPVT was eight months below that on the Binet. Burnett (1964) reported a study comparing PPVT, Wechsler-Bellevue, and S-B scores on 238 residents of a state school for the mentally retarded, ranging in age from 8 to 21 years. Correlations involving the PPVT ranged from 0.27 with Wechsler-Bellevue performance IQ's to 0.47 with Wechsler-

Bellevue verbal IQ's. It should be noted in connection with this study that the use of IQ scores on a sample of restricted IQ range and the use of Wechsler and S-B scores from existing records on the subjects may be factors in the relatively low correlations which were obtained. Kicklighter (1964) compared PPVT and S-B scores of a group of 66 educable mentally retarded children, ranging in age from 6-7 to 16-4, obtaining correlations of 0.87 for MA's and 0.71 for IQ's. In a study of 62 institutionalized retardates from 10.5 to 15.8 years of age Kimbrell (1960) found correlations between IQ's derived from the PPVT and WISC ranging from 0.30 to 0.43. Again it should be pointed out that the restricted range of IQ (40-62) may have depressed these coefficients to a considerable extent. A correlation of 0.71 between MA scores derived from PPVT and S-B on institutionalized retardates was reported by Mein (1962). Moss (1962) reported on a comparison of PPVT, S-B, and PMAT scores of 51 seven-year-old educable mentally retarded children MA scores from the PPVT correlated 0.66 with those from the S-B and 0.82 with those from the PMAT. PPVT and WISC scores of 107 retarded adults were compared by Tobias and Gorelick (1961). They obtained correlations ranging from 0.42 with WISC performance scores to 0.66 with WISC verbal scores. Overall, these results seem to indicate adequate validity for the PPVT in terms of its correlation with other measures of intelligence, though relatively low correlations and significant differences between scores have been found in some studies.

In studies utilizing measures of achievement as the criterion score, the PPVT has faired less well. Kimbrell (1960) obtained a correlation of only 0.04 between PPVT MA and grade placement as measured by the Gray-Votaw-Rodgers Achievement Test with a group of 62 institutionalized retardates. Moss (1962) obtained concurrent validity coefficients ranging from 0.32 to 0.66 and predictive validity coefficients of 0.22 to 0.43, using the Metropolitan Achievement Test and Illinois General Information Test as criteria. He also noted that both the S-B and PMAT produced higher correlations with these criteria. Utilizing the WRAT as a criterion measure, Tobias and Gorelick (1961) found a correlation of 0.52 between PPVT raw score and WRAT reading scores, and Wolfensberger (1962) found that PPVT MA's correlated 0.52 with reading, 0.33 with spelling, and 0.35 with arithmetic scores from the WRAT.

Overall, it appears that PPVT validity coefficients are consistently lower utilizing achievement criteria than using other intelligence tests as criteria. It also appears that the PPVT is less valid than the S-B, Wechsler scales, and PMAT, when tested achievement is the criteria, although in one study (Tobias & Gorelick, 1961) the PPVT was demonstrated to be superior to the WISC.

The inconsistency of available validity data on the PPVT, coupled with its strength in terms of administration time and suitability for handicapped children, seems sufficient reason for including this test in the battery evaluated in this study.

Pictorial Test of Intelligence

The PTI (French, 1963) was designed to provide an easily administered, objectively scored, individual test for assessing the intelligence of children between the ages of three and eight years. Items are divided into six subtest (Picture Vocabulary, Form Discrimination, Information and Comprehension, Similarities, Size and Number, and Immediate Recall) and are presented in order of difficulty within each test. All responses by the subject are made by simply indicating one of four symbols on a large response card which correctly depicts the answer to various stimulus questions. The test purports to measure intelligence in terms of verbal comprehension, perceptual organization, and ability to manipulate spatial and numerical symbols.

Standardization of the PTI was carried out on a total sample of 1,830 cases. The procedure used in selecting the cases assured that the standardization sample was comparable to 1960 census data as to regional area, community size, and occupational level of the father. The manual reports Kuder-Richardson reliability estimates for the full scale ranging from 0.87 to 0.93 for various age levels. Test-retest reliabilities for periods of two to six weeks yielded coefficients ranging from 0.90 to 0.96, and one study of test-retest reliability with a time lapse of between four and five years yielded a coefficient of 0.69. It should be noted that all studies involving test-retest reliability were done on rather small samples. Several validity studies, also on small Ns, are reported in the manual. In general, the coefficients obtained in these studies ranged from the high 50's to the low 80's when PTI scores were correlated with results of other intelligence tests. An example of these validity studies is one in which 32 first graders were administered the PTI, the S-B, WISC, and the Columbia Mental Maturity Scale (CMMS). Correlations involving the PTI indicated a correlation of 0.72 with the S-B, 0.65 with the WISC, and 0.53 with the CMMS. The manual lists one study of validity utilizing an achievement criterion. Rho correlations between PTI scores and teachers' ratings of children's achievement ranged from 0.38 to 0.77 with a median value of 0.53. The author notes that the size of these validity coefficients was dependent on the experiences of teachers making ratings, the size of the group being ranked, and the age level of the group. One additional study (Jenson, 1964) with a sample of 36 five- and six-year olds indicated that the PTI was equivalent to the S-B in prediction of California Achievement Test Scores over an inter-test interval of two years. Only one study, also reported in the test manual, involving retardates is available. The PTI and S-B were administered to a group of 21 MRs. A correlation of 0.44 was obtained and there was no difference between the mean IQ scores on the two instruments.

According to French (1962):

> The strengths of the PTI appear to be
>
> 1. effective discrimination among children three through eight years of age,
> 2. ease of administration,
> 3. objectivity in scoring,
> 4. usability with either physically typical or atypical children, and
> 5. the adequacy and recency of the standardization.
>
> All these aspects of the test can be viewed as strengths, though the extent to which the author has accomplished his purposes remains to be demonstrated by further research and practice. The principle weakness of the test is that it is largely untried. Additional research may suggest specific weaknesses or demonstrate other strengths. On the basis of currently available information, the only serious weakness which is apparent is the rather narrow range (CA 3-8), which limits the general utility of the instrument. Some criticism has also been made of the quality and size of the figures to which the testee must respond.

The decision to include this instrument in the study is based on its appeal in terms of apparent utility with young EMRs, its suitability for children at this level, and the careful way in which the test has been developed. Despite the paucity of reliability and validity on the test at this time, the test does show promise as an individual test of intelligence for young children and seems to deserve inclusion in the present study in an effort to obtain additional validity data.

Coloured Progressive Matrices

The CPM (Raven, 1947) was adapted from the earlier Progressive Matrices (Raven, 1938) specifically for use with children between the ages of five and 11 years and with mental retardates. The test is designed as a measure of Spearman's g factor in intelligence. The CPM consists of three sets of 12 matrices or designs each with one part removed. The subject selects the correct missing part from six or eight alternatives. The test is untimed and quite simple to administer, usually requiring about 15 minutes in examination time. The test is available in two forms which are identical except for mode of presentation, one using a board matrix with actual pieces to be filled in and the other printed replications of these matrices. Standardization of the board form was carried out on 291 children aged 5 to 10 years, and of the book form on 608 children from 5 to 11½ years of age.

Norms are presented separately for the two forms but these are extremely inadequate by American standards, yielding only percentile ranks in large steps at half year intervals.

Though the reliability and validity of the 1938 matrices have been demonstrated by extensive research, similar data for the CPM is relatively limited. The manual reports test-retest reliabilities ranging from 0.65 for children under seven to 0.90 over the full range of the test. It also reports a correlation of 0.50 with the Stanford-Binet for children less than seven years of age. Martin and Weichers (1954) compared the CPM with the WISC on a group of 100 children between nine and 10 years of age. Correlations of 0.91 with full scale IQs, 0.84 with verbal IQs and 0.83 with performance IQs were obtained. Correlations between the CPM and WISC subtest scores ranged from 0.74 to 0.47. Green and Ewert (1955) obtained a correlation of 0.78 with the Otis Intelligence Test with a group of 192 children ranging in age from 6-6 to 12-5 and correlations with California Mental Maturity Test IQs and Kuhlmann-Anderson IQs for small groups at different age levels ranging from 0.28 to 0.56. On the basis of this very limited data, it cannot be said that the validity or reliability of the test has been adequately demonstrated, but initial results suggest that the test shows some promise for general use.

Several studies are available dealing with the validity of the CPM with mentally retarded children. Stacy and Carleton (1955), reporting on a study of 150 institutionalized mental retardates, reported that the CPM correlated 0.54 with chronological age, 0.69 with S-B MA, 0.71 with S-B IQ, 0.62 with WISC full-scale scores, 0.55 with WISC full scale IQ, 0.52 with WISC performance IQ, and 0.54 with WISC verbal IQ. Correlation with WISC subtest weighted scores ranged from 0.28 to 0.48. Using a sample of 172 sub-normal adults, Stacey and Gill (1955) obtained correlations of 0.86 with the S-B, 0.68 with Wechsler-Bellevue full-scale IQ, 0.51 with Wechsler-Bellevue performance IQ and 0.56 with Wechsler-Bellevue verbal IQ. Malpass, Brown and Hake (1960) found that the CPM correlated about 0.50 with scores from various other intelligence tests.

The strength of this test lies primarily in its short administration time, simple administrative procedures, and the interest which the plates engender in most subjects. The test is also of interest as a relatively pure measure of g, though some critics (Burke, 1958) have suggested that it is not nearly so pure a measure of this factor as Raven intended. Weaknesses of the CPM include lack of sufficient reliability and validity data, lack of norms comparable to most American tests, and use of only a single type of task. In the experience of the present author, lack of sufficient floor and high dependence on the subject's level of understanding of the task might also be viewed as problems. Finally, the manual is extremely inadequate, lacking a number of characteristics listed as essential in the APA manual of technical recommendations (American Psychological Association, 1954).

Despite these weaknesses several factors suggested the inclusion of the CPM in this investigation. The fairly wide acceptance of the 1938 Matrices and fairly promising results in the few studies involving the CPM seem to warrant its use. Also, this study provides an opportunity to obtain needed validity data and to examine the factor purity of the test. Finally, along with the PPVT, the CPM provides the greatest practical advantage in terms of administration time among the individual tests being considered.

Primary Mental Abilities Test

The PMAT (Thurstone & Thurstone, 1962) is designed to provide both multi-factored and general measurement of intelligence. The 1962 revision of the test consists of five levels yielding four to five specific factor scores depending on level, as well as a general or total IQ score. The form designed for grades two to four was utilized in this study. This form of the PMAT consists of four subtests (Verbal Meaning, Spatial Relations, Number Facility, and Perceptual Speed) and yields an intelligence quotient and MA based on each factor as well as an IQ score for the total test. The test is designed for group administration, generally requiring slightly over an hour to complete, and it is suggested that this examination period be divided into two sections on successive days. The PMAT was standardized on a group of 31,919 students from kindergarten through age 12, stratified on the basis of regional location and school size.

Information concerning the statistical characteristics of the PMAT is relatively limited. The manual is quite inadequate in this respect but a technical supplement (SRA, 1965) is available which presents additional data. Because the tests are timed, estimates of internal consistency were not felt to be appropriate. Stability estimates are reported by grade level with test-retest intervals of one and four weeks. Median values over all grade levels were: total score, 0.91; Verbal Meaning, 0.89; Spatial Relations, 0.78; Number Facility, 0.81; Perceptual Speed, 0.67; and Reasoning, 0.83. Median values for level 2-4 which was used in this investigation were: total score, 0.92; Verbal Meaning, 0.86; Spatial Relations, 0.77; Number Facility, 0.88; and Perceptual Speed, 0.80. Predictive validity of PMAT IQ scores was evaluated in comparison with school grades after a time lapse of approximately 14 months. Coefficients ranged from a low of 0.03 to 0.74 for the subtests and 0.30 to 0.78 for total score. The validity for the tests at grade 2 seems most appropriate for comparison with the present study.

Median coefficients at this level were: total score, 0.53; Verbal Meaning, 0.40; Spatial Relations, 0.30; Number Facility, 0.56; and Perceptual Speed, 0.46. Correlations between PMAT scores and grades in various subject areas are reported for all grade levels combined. Coefficients for the total PMAT score were in the .50s and .60s and essentially equal for all subject areas. The PMAT was also compared with the Kuhlmann-Anderson for a group of 45 second grade children. This comparison yielded coefficients of 0.66 for the total score, 0.40 for Verbal Meaning, 0.50 for Spatial Relations, 0.65 for Number Facility, and 0.45 for Perceptual Speed. Overall, the reliability and validity of the PMAT appear to be adequate and the estimates reported in the technical supplement suggest that the reliability and validity of the tests are highest at younger age levels.

Despite these indications of reasonable reliability and validity, previous reviewers have been generally critical of the PMAT. Anastasi (1961) has criticized the PMAT on a number of counts. She suggests that multiple aptitude batteries contribute little at younger ages, that the present batteries of the PMAT have an insufficient base in research and empirical refinement and validation, and that it has inadequacies of normative data, types of scores, interpretation of scores, validity data and reliability. Kurtz (Buros, 1959) elaborated on these shortcomings in a very comprehensive review of the test. In this review Kurtz reported multiple correlations of from 0.70 to 0.80 between weighted scores on the PMAT subtests and S-B scores, and a correlation of 0.77 between PMAT quantitative score and arithmetic achievement.

Information regarding the characteristics of the PMAT with retarded children is extremely limited. Moss (1962) indicated that the test-retest reliability of the PMAT is below that of the S-B for a sample of 51 educable retardates, but he did not report these coefficients. This study yielded a correlation of 0.82 between MAs derived from the PMAT and PPVT and 0.56 between scores derived from the PMAT and the S-B. In addition, the PMAT was superior to both the Binet and PPVT in predicting reading and arithmetic achievement.

Inclusion of the PMAT in the study is based primarily on two considerations. First, because of the much greater efficiency of this type of instrument in testing within the school setting, it seems desirable to include at least one group test of intelligence in the battery. Second, available research suggests both the need for further validation studies and the promise of high validity should early estimates of validity be substantiated.

Wide Range Achievement Test

The WRAT (Jastak, 1946) is an individually administered test of achievement in reading, spelling and arithmetic from kindergarten through college level. The reading score is based on word recognition (letter recognition at the bottom of the scale). Words are listed in approximate order of difficulty and the examinee is asked to read through the list to a point where 10 consecutive words are missed. The author contends that the word recognition task is the purest measure of reading ability since the subject cannot rely on context or general knowledge for clues to identification of the word. The spelling test is administered by reading the words and having the subject write each word. At the lowest level the test requires that the subject spell his name or simply write letters of the alphabet. The arithmetic is entirely computation, again arranged in approximate order of difficulty. The author defends the use of a purely computational test on the grounds that more involved tests of mathematical ability tend to be contaminated with other ability factors, particularly reading. Examination time tends to run from 15 to 30 minutes, though group administration can be used for the spelling and arithmetic sections which may reduce total testing time for a class considerably. Raw scores can be converted directly to grade equivalent scores. These in turn can be converted to age equivalent and ratio IQ scores by means of tables.

The WRAT was standardized on approximately 4,000 students primarily in New York and Delaware. To control for the local nature of this sample, the author developed the norms so that

WRAT scores would conform to the nationally standardized New Stanford Achievement Test. Validity is reported as being checked by correlations with Stanford scores but coefficients are not presented. Test-retest reliability (reading, 0.95; arithmetic, 0.90; spelling, 0.88) are reported in the manual these figures representing data on slightly over 100 cases in each instance. The WRAT has been used as a criterion of achievement in a number of studies of the validity of intelligence tests (Wolfensberger, 1962; Tobias & Gorelick, 1961) and it has received fairly broad acceptance, especially among teachers of the mentally retarded. One study directly related to the validity of the WRAT (Lawson & Avile, 1962) demonstrated a correlation of 0.94 between reading scores on the WRAT and the Gray Oral Reading Paragraphs with a group of 30 adult retardates. A second study (Saxon, et al., 1963) compared the WRAT and the California Achievement Test performance of a group of 109 institutional retardates. Correlations of these test scores with Wechsler IQs suggested that the tests were equal in their relation to intelligence. Concurrent validity coefficients for spelling, 0.81; arithmetic, 0.83; and reading, 0.86 were reported.

In spite of the lack of adequate reliability and validity data, the WRAT has several strengths which support its inclusion in this battery. The content validity is good and the rationale underlying the reading and arithmetic tests seem particularly appropriate. It is the only individual achievement test which is generally available, the short administration time is appealing, and, perhaps most important, it appears to have sufficient floor to be used in the evaluation of young EMR children.

New York Achievement Test

The New York Achievement Tests (Wrightstone, et al., 1959) are a battery of three tests, the Wide Range Test of Reading Achievement, the Wide Range Test of Mathematical Concepts, and the New York City Core Achievement Test. The Wide Range Test of Reading Achievement is a group administered test of reading achievement providing one score for the measurement of reading skills through the fifth grade level. Skills measured include word recognition, word meanings, sentence comprehension, and paragraph comprehension. The test was designed specifically to measure reading achievement of mentally retarded children. The second test in this battery, the Wide Range Test of Mathematical Concepts, was also developed specifically for use with mentally retarded children. It is designed to measure proficiency with mathematics concepts as opposed to computational skills and includes such items as meanings of mathematical terms and symbols, money problems, telling time by clock and calendar, and miscellaneous practical problems in mathematics. The New York City Core Achievement Test is designed to measure proficiency in those elements of a typical EMR curriculum other than reading and mathematic. The Primary Core series used in this study includes questions regarding health, safety, personal care, knowledge of home and neighborhood, travel in the community, and practical purchasing.

Though these tests have not been standardized in the strictest sense, information based on administration of each test to over 1,000 mentally retarded subjects is available regarding the statistical characteristics of the test. Percentile norms based on rather small numbers of subjects are presented, but are listed as tentative. All reliability information reported for the tests is in terms of internal consistency estimates based on Froelich's approximation of a Kuder-Richardson formula and are reported separately for young, intermediate, and older groups. Coefficients are reported for the reading test ranging from 0.95 to 0.96, for mathematics from 0.81 to 0.85, and for core from 0.79 to 0.84. At the age level corresponding to that of the present study the coefficients were: reading, 0.95; mathematics, 0.82; core, 0.84. Validity of the scale is discussed only in terms of face validity of the tests as supported by teachers and supervisors in New York's Bureau CRMD.

Inclusion of the New York Achievement Test Battery in the present study is based primarily on the fact that it was developed specifically for use with mentally retarded children. It is recognized that the lack of grade level norms and lack of adequate reliability and validity data may prove a serious deficiency, but in general this seems to be one of the most appropriate group tests of achievement presently available for evaluating academic achievement in an EMR group.

Coding

Coding A for young subjects taken from the WISC (Wechsler, 1949) was included as one test requiring the subject to demonstrate learning ability within the test situation. In this test the subject is required to associate paired symbols and to write the correct response element for a random series of stimulus elements. For example, the stimulus figure "cross" must be associated with a small circle. A total of 5 stimulus figures are included, each associated with its own response. The subject has two minutes in which to complete 45 items. For the purpose of this study the raw score (total number correctly completed) and standard scores based on the experimental group's scores were utilized.

The coding test was standardized as a part of the WISC battery. The sample of 300 subjects on which this test was standardized were carefully selected to be representative of the general population (1940 census) in terms of geographic location, parental occupation, and rural-urban proportion. Reliability based on an alternate form test situation is reported as 0.60 with a standard error of estimate of 1.90 scaled score units. A test-retest reliability of 0.83 is reported (Throne, *et al.*, 1962) based on a sample of 39 retarded boys. The validity of the coding subtest has not been studied apart from the WISC as a whole, but its validity as a measure of intelligence is not an issue in the present investigation. The critical issue is whether the test provides an opportunity for the subject to demonstrate his ability to learn within the test situation. Unfortunately, it appears that two factors are important sources of variability in scoring on this test--clerical speed, and ability to learn within a very few minutes to associate certain symbols. Further, it is impossible to differentiate between the relative contribution of these two factors to the total score. However, the necessity of learning the symbol association in order to score well on the test appears sufficient to warrant its inclusion as a test of demonstrated learning ability within the context of the present study.

Modified Prognostic Reading Test

The MPRT (Woodcock, 1964) is a revision of the Prognostic Reading Test (Woodcock, 1960). It consists of three short stories written entirely with pictorial symbols instead of words. A total vocabulary of approximately 30 words is introduced in the three progressively more difficult stories. The test is administered by introducing the symbols and associating them with the correct oral response, giving each child a brief practice session, and asking him to read a story. The task of the subject is to learn to read in this new orthography within the test session and the test score is simple--the number of errors (hesitations, substitutions, omissions, and additions) which the child makes in reading the three test stories. For the purpose of the present study, these error scores were subtracted from an appropriate constant in order that scores would be positively related to performance. In addition, standard scores were derived from this distribution in order to make scores directly comparable to those from other criterion measures.

Reliability and validity data on the Modified Prognostic Reading Test are as yet unavailable. Information on the original version of the test (Woodcock, 1958) indicates concurrent validity (against Gates Reading Survey and Gates Oral Reading Test) from 0.48 to 0.65. Correlation between the PRT and the composite score on the Gates in this study was 0.59. In two studies of predictive validity of the PRT reported in the test manual, it was found that differences in scores on the PRT were related to differences in success in remedial reading. The fact that these validity studies did not include correlational techniques makes it difficult to compare the results directly with typical validity studies on test of this sort. Reliability of the test using Kuder-Richardson formula 21 is 0.91. Split-half reliability of 0.92, corrected to 0.96 by the Spearman-Brown formula, is also reported. Though information concerning the PRT is generally scant, and on the 1964 modification virtually nonexistent, the obvious need for learning within the test situation to score well on this test seems to warrant its inclusion in the criterion battery.

Paired Associate Task

The PAT consists of eight pictures of common animals to be matched with the digits from one to eight. The pairs are presented using a slide projector. The task is first explained to the child and the series of slides presented once to assure that the subject can identify each picture and name each response numeral. Testing then proceeds by presenting a slide showing a stimulus picture alone, allowing five seconds for the subject to respond, then showing the stimulus with the correct response. If the subject has responded correctly, the examiner reinforces this response verbally when the response slide appears. Five randomizations of the series are used to control for possible serial learning effects. Each subject is carried to a criterion of one correct response through a complete series, or through 10 repetitions of the series, whichever occurs first. The subject's score is the number of correct responses made, assuming that if criterion is reached before 10 trials all remaining responses would be correct, giving a total possible score of 80. This procedure differs from the usual PA task in which all subjects are tested to some pre-determined criterion and the score based on number of trials to reach that criterion. This alteration seemed justified because of the increased economy in terms of administration time and the assumption that scores would still be indicative of the number of associations learned and the facility of the subject in learning them. Directions used by examiners in the administration of the test were as follows:

> Set up projector and screen so that image fills screen as nearly as possible. Set projector for 5 second intervals. Seat subject facing screen to left of projector and fill out information on test sheet. Give the following directions.
>
> (name) , I WANT YOU TO WATCH THE SCREEN WHILE I SHOW YOU SOME SLIDES. EACH SLIDE HAS A PICTURE OF A DIFFERENT ANIMAL. FIRST WE WILL SEE EACH ANIMAL MATCHED WITH ITS NUMBER. AS I SHOW EACH SLIDE, YOU SAY THE NUMBER THAT GOES WITH EACH ANIMAL. (Show first series of slides without further comment).
>
> TRAY 1. NOW WHEN THE SLIDES ARE SHOWN THEY WILL NOT HAVE A NUMBER WITH THE PICTURE OF THE ANIMAL. WHEN YOU SEE A PICTURE OF AN ANIMAL, I WANT YOU TO TELL ME QUICKLY WHICH NUMBER GOES WITH THAT ANIMAL, AND I WILL WRITE IT DOWN. THEN THE SLIDE WILL

> COME ON WITH THE ANIMAL AND THE NUMBER, AND WE WILL SEE IF YOU GAVE ME THE RIGHT NUMBER. IF YOU DO WELL, YOU WILL GET A PRIZE WHEN WE FINISH. BE SURE AND TELL ME THE NUMBER QUICKLY BEFORE THE SLIDE WITH THE ANSWER COMES ON. (Continue through second slide series, or until subject gets every slide correct in any one series of eight. Verbally reinforce each response when the response slide appears. Correct any errors which occur in the first 16 trials. Encourage subject to respond before the second pair appears if necessary.

Research results relating to the inclusion of a paired-associate task in the battery are mixed. The number of studies (Eisman, 1958; Barnett, 1958; Wilson, 1928; Akutagawa & Benoit, 1950) suggest that this type of test fails to differentiate among children of varying MA and IQ levels. Other studies (Ellis, Pryer, & Distefano, 1960; Pryer, 1960; Sloan & Berg, 1957; McCulloch, Reswich & Roy, 1955) suggest that with more difficult material and particularly with verbal material MA differences do show up on the paired-associate task. In general, it appears, from the research to date, that the relationship between rote learning, such as the paired-associate task, and IQ or MA depends on the difficulty level of the material used and the degree of relationship between stimulus and response items. Considering the low MA of the group under consideration in the study, it was anticipated that this task should be sufficiently sensitive to differentiate along the lines of MA. In any event, the face validity of the task is good in that the associations reported are not general knowledge, but must be built up within the testing situation.

Simple Alternation Task

The Simple Alternation Task (SAT) is carried out utilizing the Yerkes Multiple Choice Apparatus. The subject's task is to press one of two levers, one of which is always correct. He must learn that to be consistently correct he must alternate between the two levers. The nature of the task if first explained to the subject and testing is begun immediately. The left hand lever is correct for all subjects on the first trial. Inter-trial interval is approximately two seconds with a 10 second rest break at the end of each block of eight trials. The test continues to a criterion of eight consecutive correct responses or a total of 80 trials, whichever occurs first. Total number of correct responses is used as the score.

Exact directions as used by examiners were as follows:

> Fill out test sheet, and sit facing the subject with YMCA between E and S. Explain task to S. IN THIS TEST YOU ARE TO FIND OUT WHICH OF TWO LEVERS IS THE RIGHT ONE TO PRESS. THE BOX HAS TWO LEVERS. ONE OF THEM WILL BE RIGHT AND ONE OF THEM WRONG EACH TIME I ASK YOU TO PRESS ONE. WHEN YOU PUSH THE RIGHT ONE YOU WILL HEAR A BUZZ (demonstrate), BUT WHEN YOU PRESS THE WRONG ONE YOU WILL HEAR NOTHING (again demonstrate). SOMETIMES ONE OF THE LEVERS WILL BE RIGHT AND SOMETIMES THE OTHER ONE WILL BE RIGHT. YOUR JOB IS TO LEARN TO PRESS THE RIGHT ONE EVERY TIME. EVERY TIME I SAY PRESS I WANT YOU TO PRESS THE LEVER THAT YOU THINK IS RIGHT FOR THAT TIME. IF YOU LEARN WELL ENOUGH TO GET THE RIGHT LEVER 8 TIMES IN A ROW YOU WILL GET A PRIZE AT THE END OF THE TEST. REMEMBER WHEN YOU PRESS THE RIGHT LEVER YOU WILL HEAR A BUZZ, BUT IF YOU PRESS THE WRONG LEVER YOU WILL HEAR NOTHING. Proceed immediately to trial one on which the lever to Ss left is correct. Alternate the correct response lever on each trial to the completion of the test. Continue testing with a minimum of

> examiner interference until the subject makes 8 consecutive correct responses or until he has gone through a total of 80 trials--whichever occurs first. Continue to encourage S to try to pick the correct lever every time. Use approximately a two-second interval with 10-second interval after each eighth trial. Directions may be repeated. Give subject the reinforcement reminding him to save it until after school.

Problem solving behavior in an alternation task has been one of the more common methods of studying problem solving in retardates. Studies by Pascal and Zak (1956) and Stolurow and Pascal (1950) found fairly high correlations, ranging from 0.64 to 0.83, between performance on a double alternating task and IQ. Denny (1964) suggests that these results indicate that the double alternation task should be a part of any new process oriented or theory-based method of assessing intelligence. He suggests that added weight is given to this recommendation by the finding that double alternation [illegible] are among those on which familials excel organics (Barnett, Ellis, & Pryor, 1960). Ringelheim (1960a, 1960b) reported low but significant correlations between MA and scores on a simple alternation learning task.

Trial testing indicated that the double alternation task was too difficult for the age and ability level represented in the experimental group so the simple alternation task was chosen. The use of alternation learning as a criterion measure is based primarily on previous research having demonstrated its relationship to mental development. On a rational basis as well, it is obvious that learning must take place within the test situation in order for the subject to succeed at this task. This, coupled with the simplicity of directions needed, and a minimal level of subject-examiner interactions, support the inclusion of this task in the battery.

APPENDIX B

INTERCORRELATIONS

104

APPENDIX B

INTERCORRELATIONS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1. Chronological Age		61	11	60	29	31	27	50	28	31	41	36	22	37	37	04	57	19	46	38	52	50	30	30	31
2. Stanford-Binet MA			85	78	52	33	35	76	38	50	41	39	44	31	57	38	72	57	65	52	66	61	61	36	49
3. Stanford-Binet IQ				59	45	22	28	63	29	43	24	26	41	16	49	49	54	60	52	39	50	44	45	25	42
4. ITPA LA					71	50	52	73	50	67	48	48	58	44	45	26	76	62	56	57	63	62	66	41	53
5. ITPA SS						38	33	43	29	55	37	31	40	17	26	12	51	43	34	39	40	37	41	28	43
6. Auditory Decoding							17	23	26	21	15	29	06	05	25	11	44	40	35	32	29	28	41	42	32
7. Visual Decoding								25	26	38	32	28	21	21	35	29	44	38	35	28	40	37	35	20	35
8. Auditory-Vocal Association									34	49	29	47	33	34	50	36	66	58	34	45	62	58	55	36	42
9. Visual Motor Association										34	01	12	07	21	15	08	46	43	31	32	41	33	47	30	32
10. Vocal Encoding											33	28	29	18	44	34	54	50	41	37	57	38	47	24	41
11. Motor Encoding												13	18	13	39	25	35	18	34	34	36	25	12	18	40
12. Auditory Vocal Automatic													10	14	50	39	39	26	40	30	21	30	33	16	32
13. Auditory Vocal Sequencing														22	00	-07	28	24	10	25	18	29	31	02	17
14. Visual Motor Sequencing															13	02	46	36	28	32	30	45	39	35	27
15. PPVT MA																89	48	38	62	27	47	32	34	25	46
16. PPVT IQ																	28	31	48	16	28	16	14	10	35
17. PTI MA																		89	72	61	78	79	82	67	68
18. PTI IQ																			64	75	68	69	72	65	64
19. Picture Vocabulary																				52	62	56	42	42	47
20. Form Discrimination																					51	49	56	53	60
21. Information & Comprehension																						59	57	43	52
22. Similarities																							65	49	57
23. Size & Number																								47	50
24. Immediate Recall																									45
25. Matrices Raw Score																									
26. Matrices St. Score																									
27. PMA Raw Score																									
28. PMA MA																									
29. Verbal Meaning																									
30. Spatial Relations																									
31. Number Facility																									
32. Perceptual Speed																									
33. WRAT Reading AE																									
34. WRAT Reading SS																									
35. WRAT Spelling AE																									
36. WRAT Spelling SS																									
37. WRAT Arithmetic AE																									
38. WRAT Arithmetic SS																									
39. NY Reading RS																									
40. NY Reading SS																									
41. NY Math RS																									
42. NY Math SS																									
43. NY Core RS																									
44. NY Core SS																									
45. CRITERION I																									
46. Coding RS																									
47. Coding SS																									
48. MPRT RS																									
49. MPRT SS																									
50. Paired Associates SS																									
53. Simple Alternation SS																									
54. CRITERION II																									
55. CRITERION III																									

APPENDIX B

INTERCORRELATIONS

20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	
28	52	50	10	30	31	31	58	51	49	33	52	45	32	32	42	42	52	52	43	43	47	47	41	41	55	44	44	43	43	42	42	24	24	56	59	1
52	66	61	61	36	49	49	70	61	58	43	61	34	28	28	52	52	71	71	53	53	66	66	67	67	73	59	59	41	41	53	53	36	35	69	76	2
39	50	44	45	25	42	42	50	43	41	33	42	38	15	14	39	38	55	55	37	37	51	52	57	57	55	46	47	26	28	39	39	30	30	51	57	3
37	63	62	46	41	53	53	66	57	53	46	61	47	26	26	46	46	67	67	50	50	61	61	51	51	63	47	47	48	48	41	41	34	34	62	69	4
39	40	37	41	26	43	43	48	40	35	37	52	29	21	21	36	36	41	41	32	32	40	40	35	35	46	27	27	30	29	10	11	20	21	32	43	5
32	29	28	41	42	32	33	34	29	39	15	31	20	00	00	13	13	28	28	22	21	25	26	29	29	25	26	27	09	09	15	16	14	15	25	27	6
28	40	37	35	20	35	34	23	20	18	27	22	09	00	01	02	02	26	26	21	21	24	24	19	19	20	14	14	13	13	19	19	09	09	20	22	7
45	62	58	55	36	42	42	66	59	63	37	55	49	27	27	49	49	65	65	44	44	59	59	60	60	65	47	47	48	49	53	53	27	27	64	69	8
32	41	33	47	30	32	32	45	46	29	41	38	36	07	07	21	21	46	45	31	32	39	39	28	28	36	39	38	23	24	12	12	23	23	35	39	9
37	57	38	47	24	41	41	39	31	35	37	30	25	07	07	27	27	37	36	26	26	33	33	44	44	38	30	30	21	21	22	22	21	21	35	39	10
34	36	35	12	18	40	40	28	25	27	29	23	16	07	07	10	10	15	15	19	19	17	17	29	29	21	17	17	21	22	19	19	18	18	28	28	11
30	21	30	33	16	32	32	35	27	45	13	34	14	12	12	19	18	34	34	37	36	42	42	38	38	39	21	22	23	23	38	38	28	28	39	42	12
25	18	29	31	02	17	17	27	21	11	15	26	28	25	25	33	33	40	40	25	25	39	39	14	14	38	19	20	39	39	18	17	24	25	37	41	13
32	30	45	39	35	27	26	41	38	22	27	44	34	22	22	39	39	46	47	38	39	37	37	20	20	43	25	26	33	32	41	41	04	04	37	43	14
27	47	32	34	25	46	46	42	37	34	34	26	24	02	02	12	13	27	27	23	22	31	31	30	30	32	39	39	20	20	39	39	35	34	48	42	15
16	28	16	14	10	35	35	20	16	35	18	05	07	-07	-07	03	02	10	10	06	05	13	13	35	35	12	23	23	11	10	21	21	32	31	31	22	16
81	78	79	62	67	68	67	78	70	60	59	69	59	27	27	48	48	75	75	55	56	59	59	39	39	70	57	57	42	42	44	44	28	28	62	71	17
75	68	69	72	65	64	64	64	38	47	53	53	49	15	14	26	26	63	63	44	44	47	47	48	49	34	45	45	28	29	32	31	20	19	44	34	18
62	62	36	42	42	47	47	58	53	56	46	44	38	17	16	35	35	49	49	41	40	42	42	38	38	52	50	50	24	24	42	42	28	25	51	55	19
	51	49	38	53	60	60	38	49	36	43	51	51	23	23	36	36	56	56	51	51	38	38	41	42	33	46	46	33	33	30	30	22	22	45	53	20
		39	37	43	52	53	39	53	51	47	52	38	17	17	32	33	61	61	33	33	47	48	48	48	51	43	43	40	40	40	40	14	13	30	34	21
			65	49	57	57	63	61	56	49	52	45	23	22	41	41	61	61	42	42	54	53	53	53	59	43	44	23	23	46	46	29	29	32	59	22
				47	50	50	65	59	47	44	61	53	27	27	48	48	63	63	48	48	57	57	48	48	62	48	49	38	38	41	41	18	18	54	62	23
					45	45	52	47	39	43	42	42	11	11	23	23	53	53	33	32	35	35	35	36	41	45	46	20	20	30	30	18	18	42	43	24
						100	59	55	42	36	50	42	08	08	38	38	48	48	43	43	45	45	31	51	48	47	47	21	34	32	32	32	32	48	30	25
							59	55	42	36	40	41	08	07	28	28	48	48	42	42	45	45	30	51	47	47	47	23	24	27	32	32	32	48	50	26
								96	76	72	84	84	31	30	55	55	77	76	66	66	73	73	65	65	79	63	63	48	48	44	43	30	30	67	79	27
									73	79	80	79	26	25	50	51	69	69	63	62	70	70	61	61	73	58	58	41	41	38	38	27	27	60	71	28
										42	52	51	13	12	28	27	52	52	43	43	62	63	66	66	57	50	50	31	32	42	42	27	27	35	60	29
											56	52	07	07	32	33	47	47	48	47	46	46	44	44	48	34	35	32	32	28	25	30	30	41	48	30
												37	29	29	61	61	76	76	65	65	71	71	48	48	77	50	50	50	50	38	38	16	16	35	71	31
													29	29	48	48	62	63	34	53	52	52	47	47	62	60	60	41	41	32	32	23	22	37	64	32
														100	80	80	48	49	46	47	26	26	16	16	68	22	22	49	49	32	32	09	09	41	60	33
															79	79	48	49	46	46	26	25	15	16	68	22	22	49	49	32	32	09	09	41	59	34
																100	68	68	60	61	49	49	38	38	85	66	66	57	57	46	46	16	16	60	79	35
																	68	68	61	61	50	49	38	38	85	66	66	57	57	46	46	16	16	60	80	36
																		100	66	66	70	70	50	50	87	60	60	34	34	53	53	27	27	71	83	37
																			66	66	70	70	50	50	87	60	60	53	53	53	53	27	27	70	85	38
																				100	57	57	47	47	81	41	41	42	42	44	44	13	14	54	73	39
																					57	58	47	47	81	41	41	42	42	44	44	14	14	51	73	40
																						100	58	58	79	69	69	43	43	49	49	28	28	61	75	41
																							58	58	78	69	69	43	43	49	49	28	28	61	75	42
																								100	68	50	50	23	24	46	46	28	28	54	65	43
																									67	50	51	24	24	46	46	28	28	54	65	44
																										56	56	58	58	58	58	28	28	73	84	45
																											100	34	35	60	60	30	30	75	79	46
																												33	33	60	60	31	30	75	71	47
																													100	33	33	15	15	67	67	48
																														33	33	15	15	67	67	49
																															100	21	21	70	68	50
																																21	21	70	69	51
																																	100	61	66	52
																																		61	66	53
																																			63	54

APPENDIX C

RAW TEST DATA

APPENDIX C

RAW TEST DATA

Variable/Subject	65	71
1. Chronological Age	120	119
2. Stanford-Binet MA	98	84
3. Stanford-Binet IQ	80	69
4. Illinois Test of Psycholinguistic Abilities LA	95	73
5. Illinois Test of Psycholinguistic Abilities SS	443	209
6. ITPA Auditory Decoding LA	106	95
7. ITPA Visual Decoding LA	70	75
8. ITPA Auditory-Vocal Association LA	87	82
9. ITPA Visual-Motor Association LA	90	78
10. ITPA Vocal Encoding LA	83	79
11. ITPA Motor Encoding LA	70	42
12. ITPA Auditory-Vocal Automatic LA	78	69
13. ITPA Auditory-Vocal Sequencing LA	102	71
14. ITPA Visual-Motor Sequencing LA	88	68
15. Peabody Picture Vocabulary Test MA	66	59
16. Peabody Picture Vocabulary Test IQ	61	55
17. Pictorial Test of Intelligence MA	99	92
18. Pictorial Test of Intelligence IQ	80	75
19. PTI Picture Vocabulary MA	90	84
20. PTI Form Discrimination MA	108	108
21. PTI Information & Comprehension MA	84	90
22. PTI Similarities MA	87	72
23. PTI Size & Number MA	90	108
24. PTI Immediate Recall MA	108	75
25. Coloured Progressive Matrices RS	17	14
26. Coloured Progressive Matrices SS	58	52
27. Primary Mental Abilities RS	122	100
28. Primary Mental Abilities MA	98	92
29. PMA Verbal Meaning MA	92	82
30. PMA Spatial Relations MA	88	82
31. PMA Number Facility MA	104	102
32. PMA Perceptual Speed MA	126	114
33. Wide Range Achievement Test: Reading RS	89	91
34. Wide Range Achievement Test: Reading SS	64	66
35. Wide Range Achievement Test: Spelling RS	86	91
36. Wide Range Achievement Test: Spelling SS	64	72
37. Wide Range Achievement Test: Arithmetic RS	97	95
38. Wide Range Achievement Test: Arithmetic SS	71	68
39. New York: Reading RS	44	58
40. New York: Reading SS	66	77
41. New York: Mathematics RS	18	18
42. New York: Mathematics SS	68	68
43. New York: Core RS	34	31
44. New York: Core SS	64	59

RAW TEST DATA (continued)

Variable/Subject	65	71
45. Criterion I	66	68
46. Coding RS	43	45
47. Coding SS	62	63
48. Modified Prognostic Reading Test RS	115	117
49. Modified Prognostic Reading Test SS	60	61
50. Paired Associate Task RS	41	69
51. Paired Associate Task SS	48	62
52. Simple Alternation Task RS	80	40
53. Simple Alternation Task SS	57	36
54. Criterion II	57	56
55. Criterion III	62	62

RAW TEST DATA (continued)

Var./S	95	41	62	85	77	88	64	48	1	91	79
1	122	110	107	117	109	144	122	118	119	109	122
2	82	84	80	96	84	80	88	84	88	82	78
3	67	74	73	80	75	69	71	70	73	73	63
4	84	81	75	84	67	66	79	74	78	76	68
5	341	305	233	335	231	200	287	227	269	251	200
6	106	53	53	65	77	74	106	65	77	106	95
7	87	105	70	105	80	44	105	80	80	80	80
8	70	92	87	82	78	70	73	70	99	82	66
9	103	82	73	86	99	61	78	90	90	69	61
10	76	68	83	107	57	61	107	53	68	46	53
11	95	82	70	88	76	76	76	55	104	42	60
12	78	69	82	51	96	55	60	46	64	96	87
13	84	102	75	102	30	102	94	102	71	102	52
14	68	72	108	64	64	52	50	72	68	76	72
15	78	76	68	87	96	63	84	64	76	71	91
16	71	77	70	77	91	58	69	59	69	73	80
17	108	75	100	83	84	68	94	90	83	92	78
18	87	66	91	70	74	58	76	75	69	82	63
19	108	84	102	78	108	78	108	72	102	93	102
20	108	66	108	72	102	51	75	102	78	108	93
21	102	66	72	78	78	66	102	63	84	63	60
22	102	84	96	78	69	78	78	108	78	87	60
23	108	78	108	108	75	69	90	108	84	108	78
24	102	66	96	75	75	66	96	96	96	87	102
25	21	23	17	18	17	11	19	19	18	16	12
26	68	72	58	61	58	45	63	63	61	56	47
27	108	92	120	90	79	85	91	97	109	86	63
28	99	96	102	95	94	95	88	93	103	86	80
29	92	82	98	76	92	100	86	82	86	92	78
30	106	108	112	128	96	94	88	102	102	80	82
31	106	94	100	94	78	80	88	84	104	100	78
32	92	98	114	96	98	96	114	126	108	78	98
33	92	82	86	74	78	80	82	86	83	78	82
34	67	56	60	47	51	53	56	60	57	51	56
35	89	89	90	83	78	82	83	83	83	82	82
36	69	69	70	60	52	58	60	60	60	58	58
37	103	86	100	86	86	80	90	90	88	92	82
38	78	58	74	58	58	50	62	62	60	65	53
39	46	34	63	41	66	32	32	44	29	34	36
40	68	59	81	64	83	57	57	66	55	59	60
41	19	17	17	21	14	17	8	15	16	19	13
42	70	66	66	75	59	66	45	61	63	70	56
43	28	33	30	27	40	38	35	28	36	26	27
44	54	62	57	52	74	70	65	54	67	50	52

RAW TEST DATA (continued)

Var./S	95	41	62	85	77	88	64	48	1	91	79
45	68	62	68	59	63	59	58	60	60	59	56
46	42	45	41	45	45	44	44	46	34	28	41
47	61	63	60	63	63	62	62	64	54	50	60
48	118	123	107	101	92	121	118	120	116	125	115
49	62	64	57	54	50	63	62	63	61	65	60
50	28	61	76	78	72	73	66	46	61	60	70
51	41	58	65	66	63	64	60	50	58	57	62
52	76	80	40	79	79	78	80	78	76	80	80
53	55	57	36	56	56	56	57	56	55	57	57
54	55	60	54	60	58	61	60	58	57	57	60
55	62	61	61	60	60	60	59	59	58	58	58

Var./S	31	54	21	87	69	99	80	47	42	26	55
1	105	119	120	120	113	116	105	108	110	108	110
2	76	66	96	83	78	78	74	72	72	68	66
3	70	55	78	68	67	66	68	65	64	61	58
4	76	70	74	73	72	81	59	66	71	66	65
5	251	200	227	209	200	311	200	200	200	200	200
6	68	74	71	90	68	106	55	55	53	68	65
7	87	57	66	70	87	94	87	80	80	66	66
8	78	78	82	53	63	66	50	66	70	66	87
9	78	44	78	73	103	61	56	61	69	78	78
10	107	79	61	57	72	76	64	49	72	61	61
11	104	50	104	82	60	88	50	46	55	76	42
12	78	78	33	33	46	73	64	73	42	55	64
13	55	84	67	102	64	102	52	102	102	84	52
14	64	88	81	81	108	72	61	68	72	52	76
15	112	68	82	63	68	57	78	71	44	70	64
16	102	63	74	58	70	55	78	73	47	72	67
17	101	74	83	78	99	97	71	64	68	68	89
18	94	61	68	64	85	82	65	57	60	61	78
19	108	69	96	84	72	102	102	66	60	72	84
20	108	57	84	72	96	102	45	54	72	66	84
21	102	66	96	72	90	66	60	66	60	66	72
22	78	72	90	90	99	108	99	60	66	60	99
23	96	84	78	84	108	102	72	78	66	78	96
24	102	108	78	66	102	108	87	45	84	66	102
25	29	13	18	17	19	19	8	13	8	11	21
26	86	49	61	58	63	63	38	49	38	45	68
27	92	87	96	90	102	58	48	57	47	72	78
28	91	92	95	98	105	81	77	77	76	84	89
29	88	90	90	80	76	76	82	76	72	84	88

RAW TEST DATA (continued)

Var./S	31	54	21	87	69	99	80	47	42	26	55
30	124	94	96	112	116	94	84	78	74	102	90
31	86	82	94	88	100	76	72	74	78	82	80
32	94	110	102	108	110	86	70	100	82	84	98
33	80	79	82	83	80	82	85	86	95	82	85
34	53	52	56	57	53	56	59	60	71	56	59
35	79	78	79	80	83	85	83	83	88	78	85
36	54	52	54	55	60	63	60	60	67	52	63
37	82	84	88	92	95	80	86	88	86	80	84
38	53	55	60	65	68	50	58	60	58	50	55
39	35	28	35	40	22	39	24	24	43	37	38
40	60	54	60	63	50	63	51	51	66	61	62
41	11	16	12	11	16	10	11	12	16	8	13
42	52	63	54	52	63	50	52	54	63	45	56
43	33	31	36	27	33	32	32	24	20	32	33
44	62	59	67	52	62	60	60	47	40	60	62
45	56	56	58	57	59	57	57	55	61	54	60
46	33	38	34	40	28	22	35	22	20	28	37
47	54	58	54	59	50	45	55	45	43	50	57
48	118	123	107	89	108	102	80	123	100	117	84
49	62	64	57	49	57	55	45	64	54	61	47
50	64	50	44	63	52	69	68	68	66	72	38
51	59	52	49	59	53	62	61	61	60	63	46
52	79	79	75	79	77	80	79	78	66	78	80
53	56	56	54	56	55	57	56	56	50	56	57
54	58	58	54	56	54	55	54	56	52	58	52
55	57	57	56	56	56	56	56	56	56	56	56

Var./S	29	70	39	22	11	14	46	8	12	32	50
1	115	108	100	124	124	103	102	116	119	117	104
2	80	74	72	96	86	84	77	76	74	73	72
3	68	66	70	76	69	79	73	64	61	61	67
4	75	67	66	67	74	71	64	75	72	65	64
5	239	200	200	200	221	200	200	233	200	200	200
6	68	57	62	53	53	74	60	106	77	62	55
7	87	105	66	80	75	75	87	87	70	80	66
8	87	78	78	92	78	87	59	66	73	82	66
9	82	78	82	65	82	90	48	86	78	48	73
10	53	53	57	57	64	83	49	68	57	57	76
11	76	55	60	70	104	50	60	70	104	60	42
12	100	73	46	69	55	91	55	69	69	51	37
13	84	102	88	67	102	47	102	67	61	61	94
14	61	76	61	68	76	58	58	72	72	94	68

RAW TEST DATA (continued)

Var./S	29	70	39	22	11	14	46	8	12	32	50
15	78	68	57	82	70	87	76	82	57	71	54
16	71	70	60	74	64	85	77	74	53	66	57
17	86	94	81	74	72	84	74	88	78	76	66
18	73	84	86	60	58	79	70	74	65	64	61
19	108	63	96	102	66	84	84	96	66	93	66
20	66	102	84	54	60	72	54	72	84	54	78
21	78	72	78	78	60	84	78	78	78	84	57
22	96	96	108	90	90	87	84	108	78	72	51
23	90	108	66	78	72	108	84	90	72	84	78
24	75	108	96	66	96	63	66	102	108	75	63
25	17	19	17	10	15	15	13	20	12	10	11
26	58	63	56	42	54	54	49	65	47	42	45
27	94	70	84	62	61	62	61	54	81	67	70
28	98	82	97	82	81	79	83	79	83	82	84
29	86	80	74	86	80	92	76	80	84	76	72
30	106	72	108	82	82	78	88	88	82	82	88
31	90	84	86	78	80	80	82	78	78	82	80
32	108	100	112	82	86	74	84	74	120	98	114
33	77	76	80	77	79	79	92	76	74	85	82
34	50	49	53	50	52	52	67	49	47	59	56
35	76	72	83	79	79	80	82	76	74	83	84
36	49	43	60	54	54	55	58	49	46	60	61
37	85	92	86	86	86	84	80	82	86	82	84
38	56	65	58	58	58	55	50	53	58	53	55
39	19	29	36	17	26	15	24	11	19	41	19
40	48	55	60	46	53	45	51	42	48	64	48
41	18	12	15	10	16	17	10	14	12	12	10
42	66	54	61	50	63	66	50	59	54	54	50
43	29	25	29	31	26	32	21	30	29	29	24
44	55	49	55	59	50	60	42	57	55	55	47
45	54	52	58	53	55	56	53	52	51	58	53
46	30	45	39	37	28	24	43	40	38	18	47
47	51	63	58	57	50	46	62	59	58	41	65
48	06	118	89	90	104	92	81	93	89	126	105
49	56	62	49	49	56	50	45	51	49	65	56
50	68	47	38	74	52	72	[illegible]	67	75	77	44
51	61	51	46	64	53	63	61	61	65	66	49
52	79	76	79	79	63	68	79	79	75	33	66
53	56	55	56	56	48	51	56	56	54	38	50
54	56	58	52	56	52	52	56	57	56	51	55
55	55	55	55	54	54	54	54	54	54	54	54

RAW TEST DATA (continued)

Var./S	59	17	23	18	19	20	96	66	74	9	90
1	113	113	110	113	99	123	103	108	114	116	111
2	82	86	82	80	78	76	73	72	70	68	66
3	71	74	72	69	76	62	69	65	60	58	58
4	70	69	67	77	72	58	58	60	63	72	78
5	200	200	200	263	301	200	200	200	200	200	275
6	85	53	55	85	74	53	65	57	65	81	81
7	80	75	75	94	66	66	44	80	57	94	75
8	82	73	73	82	70	59	78	59	59	66	70
9	78	94	61	73	82	61	44	52	94	94	82
10	64	79	79	107	88	57	42	46	79	53	107
11	46	76	76	104	50	60	55	50	46	60	82
12	69	46	37	60	60	73	28	69	33	64	78
13	58	71	102	55	102	45	67	67	58	75	67
14	76	64	58	64	58	55	76	61	64	72	72
15	84	76	64	86	59	71	68	68	76	63	74
16	83	77	67	86	69	68	70	70	69	58	75
17	71	84	77	92	67	57	70	67	82	70	74
18	61	73	68	79	69	47	65	60	70	59	65
19	108	108	84	72	66	60	66	51	84	63	90
20	45	75	93	108	72	48	72	75	75	72	94
21	84	84	84	90	66	60	66	66	66	72	78
22	99	72	60	87	54	54	78	69	72	66	69
23	60	96	75	102	78	72	72	66	96	75	62
24	75	75	75	87	54	54	66	87	108	84	84
25	8	12	17	20	9	12	11	19	10	12	15
26	38	47	58	65	40	47	45	63	42	47	54
27	64	59	56	84	43	52	52	37	58	50	52
28	83	80	80	89	73	78	81	76	82	77	78
29	98	86	76	88	74	72	72	78	76	80	82
30	66	74	100	90	72	76	88	76	88	88	84
31	76	74	74	80	78	80	74	76	80	74	74
32	90	90	84	112	72	88	90	60	84	72	74
33	74	73	72	74	83	101	77	71	82	74	74
34	47	45	44	47	57	78	50	43	56	47	47
35	78	72	80	72	84	90	82	72	85	71	76
36	52	43	55	43	61	70	58	43	63	42	49
37	79	82	80	83	80	78	82	79	82	78	80
38	49	53	50	54	50	48	53	49	53	48	50
39	14	10	27	28	18	37	9	12	30	20	10
40	44	41	54	54	47	61	40	42	56	48	41
41	11	15	11	11	9	8	4	13	6	9	9
42	52	61	52	52	47	45	36	56	40	47	47
43	29	28	34	32	29	23	26	29	27	22	25
44	55	54	64	60	55	45	50	55	52	44	49
45	50	50	53	52	53	58	48	48	53	46	47

RAW TEST DATA (continued)

Var./S	59	17	23	18	19	20	96	66	74	9	90
46	43	45	28	44	27	27	28	45	35	33	45
47	62	63	50	62	49	49	50	63	55	54	63
48	71	84	100	78	106	101	104	98	89	109	111
49	41	47	54	44	56	54	56	53	49	58	59
50	77	58	28	47	25	35	72	44	73	74	40
51	66	56	41	51	40	45	63	49	64	64	47
52	78	63	79	79	80	40	66	80	47	79	75
53	56	48	56	56	57	36	50	57	40	56	54
54	56	54	50	53	50	46	55	56	52	58	56
55	53	52	52	52	52	52	52	52	52	52	52

Var./S	33	82	7	13	56	38	92	2	49	81	63
1	105	112	99	103	108	100	110	109	108	108	116
2	66	58	82	82	81	76	74	68	66	63	62
3	61	51	80	77	73	74	65	61	59	56	53
4	72	53	71	78	62	72	63	56	54	50	58
5	203	200	200	275	200	200	200	200	200	200	200
6	106	53	77	106	71	53	53	57	51	60	65
7	87	57	87	80	70	80	87	53	57	66	66
8	82	53	63	78	66	70	78	66	59	53	59
9	103	56	90	90	69	90	82	56	73	78	65
10	64	38	76	79	46	61	61	53	38	49	57
11	42	70	46	50	95	55	65	70	46	42	46
12	64	37	73	55	33	42	33	33	28	28	28
13	61	47	75	102	61	102	64	52	71	30	50
14	61	61	61	52	52	76	58	58	52	64	81
15	82	54	66	68	82	57	59	46	57	57	61
16	81	57	76	70	81	67	62	44	60	60	56
17	74	65	80	79	70	70	75	64	70	62	69
18	82	54	84	82	62	73	66	57	62	56	58
19	72	66	102	72	93	66	66	48	69	66	78
20	75	72	84	45	78	78	72	54	75	48	45
21	66	54	72	96	72	72	102	84	60	54	84
22	60	54	72	99	42	78	72	60	60	78	78
23	84	66	75	108	69	66	72	72	69	64	69
24	96	72	96	75	87	54	90	60	102	72	63
25	12	7	13	14	13	13	15	8	10	12	13
26	47	35	49	52	49	49	54	38	42	47	49
27	77	53	67	65	46	50	49	33	72	108	57
28	92	81	84	87	76	83	77	74	88	107	79
29	88	72	74	92	80	68	74	72	76	98	78
30	96	80	102	90	72	90	82	76	100	116	80

RAW TEST DATA (continued)

Var./S	33	82	7	13	56	38	92	2	49	81	63
31	78	82	86	72	74	76	76	78	76	82	88
32	98	82	82	84	78	90	78	60	112	122	68
33	82	107	66	73	73	80	77	80	77	64	79
34	56	85	37	45	45	53	50	53	50	35	52
35	74	90	68	72	71	83	72	82	76	71	79
36	46	70	37	43	42	60	43	58	49	42	54
37	80	68	85	80	73	86	86	79	90	73	88
38	50	36	56	50	42	58	58	49	62	42	60
39	32	0	40	8	28	36	11	11	17	24	27
40	57	33	63	39	54	60	42	42	46	51	54
41	8	6	14	18	10	8	11	10	9	16	11
42	45	40	59	68	50	45	52	50	47	63	52
43	22	22	35	32	31	18	26	18	22	33	23
44	44	44	65	60	59	37	50	37	44	62	45
45	50	51	53	51	49	52	49	48	50	49	53
46	20	33	25	22	44	15	12	18	35	28	40
47	43	54	47	45	62	39	37	41	55	50	59
48	120	121	71	68	102	110	104	102	87	82	86
49	63	63	41	39	55	58	56	55	48	46	47
50	35	17	36	57	33	38	52	65	42	49	42
51	45	36	45	56	44	46	53	60	48	52	48
52	80	67	79	78	48	59	80	64	72	71	48
53	57	50	56	56	40	46	57	49	53	52	40
54	52	51	47	49	50	47	51	51	51	50	48
55	51	51	50	50	50	50	50	50	50	50	50

Var./S	27	53	36	97	24	6	45	93	76	57	83
1	100	100	108	103	105	113	106	119	118	102	96
2	55	76	74	68	82	80	76	74	68	65	64
3	53	74	66	64	76	69	70	61	57	61	64
4	54	67	65	61	81	65	74	66	66	70	61
5	200	200	200	200	311	200	227	200	200	200	200
6	29	60	71	68	65	53	81	74	74	68	57
7	70	105	105	70	105	75	70	57	75	80	70
8	39	70	50	56	78	73	78	73	63	56	59
9	61	35	90	90	78	86	48	48	86	73	78
10	42	64	57	53	105	57	68	64	61	76	49
11	70	82	76	50	95	50	95	76	60	70	55
12	42	55	42	33	33	55	46	69	37	69	46
13	84	84	58	58	88	71	102	71	64	79	52
14	72	68	58	64	76	68	61	64	68	68	88
15	55	82	66	61	78	71	74	57	87	59	57

RAW TEST DATA (continued)

Var./S	27	53	36	97	24	6	45	93	76	57	83
16	66	90	68	64	78	73	75	55	77	62	67
17	38	62	82	62	102	90	69	80	84	64	61
18	35	60	74	58	95	78	63	66	70	60	60
19	60	66	102	63	108	96	60	102	102	54	66
20	42	54	108	45	102	108	72	84	72	72	48
21	36	54	72	54	102	78	60	72	96	63	42
22	36	66	84	51	99	78	72	87	72	60	72
23	30	66	72	90	90	78	72	72	84	62	66
24	30	60	75	75	108	102	78	75	96	96	72
25	7	10	20	9	16	10	13	15	11	11	8
26	35	42	65	40	56	42	49	54	45	45	38
27	13	32	70	51	69	70	60	53	64	28	43
28	66	70	87	78	83	87	80	80	85	71	76
29	66	74	82	66	82	80	86	80	82	66	68
30	68	68	106	82	120	94	74	82	94	78	78
31	66	72	76	74	76	80	70	80	76	74	74
32	60	66	92	106	84	98	100	74	88	68	78
33	107	77	79	77	78	73	79	76	72	72	77
34	85	50	52	50	51	45	52	49	44	44	50
35	79	72	76	79	78	71	78	72	72	71	76
36	54	43	49	54	52	42	52	43	43	42	49
37	74	77	78	79	80	80	79	78	78	79	68
38	43	47	48	49	50	50	49	48	48	49	36
39	26	10	26	13	12	17	19	20	10	16	26
40	53	41	53	43	42	46	48	48	41	45	53
41	6	9	11	10	5	10	6	6	8	8	7
42	40	47	52	50	38	50	40	40	45	45	43
43	15	31	31	21	30	28	20	27	21	23	23
44	32	59	59	42	57	54	40	52	42	45	45
45	51	48	52	48	48	48	47	47	44	45	46
46	4	15	34	35	17	45	31	28	34	29	13
47	30	39	54	55	41	63	52	50	54	50	37
48	103	85	53	95	88	80	111	91	108	70	101
49	55	47	33	51	48	45	59	50	57	40	54
50	49	65	27	32	29	36	37	48	41	73	56
51	52	60	41	43	42	45	46	51	48	64	55
52	80	75	80	68	80	55	43	69	70	76	79
53	57	54	57	51	57	44	38	51	52	55	56
54	48	50	46	50	47	49	49	50	53	52	50
55	50	49	49	49	48	48	48	48	48	48	48

RAW TEST DATA (continued)

Var./S	28	37	10	61	89	98	73	40	16	67	58
1	103	109	103	102	116	96	102	110	114	89	84
2	62	72	69	67	64	56	71	63	74	63	61
3	58	64	65	63	54	56	67	56	64	68	70
4	55	59	55	60	67	58	69	65	77	56	55
5	200	200	200	200	200	200	200	200	257	200	205
6	106	85	46	95	65	57	95	85	55	62	60
7	57	66	53	57	80	57	87	94	105	57	70
8	47	66	42	42	53	59	66	47	73	56	53
9	78	61	56	61	99	99	99	69	99	90	73
10	46	64	49	46	68	57	61	64	93	42	57
11	50	70	55	76	76	46	50	70	46	46	50
12	28	46	33	60	55	33	28	33	64	46	37
13	50	35	94	67	47	47	55	71	102	50	61
14	40	50	68	40	88	68	93	58	68	52	31
15	44	84	54	82	59	44	68	47	71	61	61
16	47	83	57	81	55	55	70	55	66	77	77
17	69	74	69	56	80	66	70	62	72	51	57
18	70	66	70	53	67	69	72	55	62	54	64
19	63	72	66	78	72	78	84	60	78	60	66
20	78	72	102	51	66	72	54	60	51	51	42
21	45	60	60	48	72	63	60	48	78	44	48
22	72	72	86	33	90	60	60	66	69	39	54
23	84	78	72	62	108	60	75	72	96	60	62
24	96	108	54	63	90	72	108	42	48	45	72
25	16	14	16	13	14	14	12	9	13	13	10
26	56	52	56	49	52	52	47	40	49	49	42
27	48	83	48	52	54	40	45	33	70	32	45
28	81	92	70	79	80	74	76	71	88	74	75
29	70	96	70	80	74	74	76	74	86	66	76
30	88	88	66	78	88	72	84	68	90	80	74
31	76	82	70	74	80	74	72	74	78	70	72
32	86	98	114	84	80	76	76	66	90	78	84
33	82	71	76	68	64	79	71	73	68	60	73
34	56	43	49	40	35	52	43	45	40	30	45
35	79	72	72	71	70	76	72	72	72	72	68
36	54	43	43	42	40	49	43	43	43	43	37
37	79	78	72	74	73	79	82	73	79	72	72
38	49	48	41	43	42	49	53	42	49	41	41
39	18	18	10	12	29	9	10	20	10	8	15
40	47	47	41	42	55	40	41	48	41	39	45
41	7	12	5	11	7	10	7	10	10	4	5
42	43	54	38	52	43	50	43	50	50	36	38
43	17	31	22	29	20	22	21	25	24	23	22
44	35	59	44	55	40	44	42	49	47	45	44
45	47	49	43	46	42	47	44	46	45	39	42

RAW TEST DATA (continued)

Var./S	28	37	10	61	89	98	73	40	16	67	58
46	34	25	31	22	28	20	42	12	31	46	28
47	54	47	52	45	50	43	61	37	52	64	50
48	51	60	92	94	82	122	43	89	105	86	91
49	32	36	50	51	46	64	28	49	56	47	50
50	40	30	39	25	59	35	44	29	20	19	21
51	47	42	47	40	57	45	49	42	37	37	38
52	80	79	79	74	78	40	78	76	38	79	77
53	57	56	56	54	56	36	56	55	35	56	55
54	48	45	51	48	52	47	48	46	45	51	48
55	48	47	47	47	47	47	46	46	45	45	45

Var./S	34	78	72	86	94	60	68	25	75	15	44
1	107	96	111	90	111	88	97	100	93	113	103
2	70	66	66	64	63	60	59	70	68	66	59
3	63	66	58	68	55	65	58	68	71	57	55
4	61	54	63	53	67	49	52	63	58	57	67
5	200	200	200	200	200	200	200	200	200	200	200
6	65	51	85	55	62	53	60	85	106	51	74
7	75	66	57	62	94	62	53	66	70	80	87
8	56	59	59	50	53	50	39	63	56	53	63
9	86	86	82	49	56	48	86	48	52	78	86
10	68	61	53	72	76	64	42	68	42	88	53
11	46	35	46	55	88	50	33	76	60	55	70
12	78	33	64	60	51	28	42	55	42	55	46
13	41	52	64	88	75	39	58	58	37	31	71
14	47	47	52	88	64	55	55	50	61	50	64
15	68	54	54	84	55	68	42	74	63	74	44
16	70	64	64	92	59	85	40	83	73	75	51
17	65	70	60	62	73	57	49	67	66	56	63
18	54	74	53	65	64	62	48	68	69	49	59
19	63	72	51	72	72	60	42	108	72	63	60
20	60	66	51	57	84	48	36	78	60	42	66
21	63	84	60	57	72	57	46	60	78	63	54
22	60	66	60	60	66	69	42	60	51	66	66
23	72	69	63	60	78	50	62	50	63	54	66
24	63	63	90	78	46	72	63	75	90	45	54
25	14	9	13	17	12	21	10	14	14	10	9
26	52	40	49	58	47	68	42	52	52	42	40
27	45	40	40	42	43	34	52	42	45	35	43
28	75	74	73	74	78	72	81	75	76	73	76
29	72	70	72	72	70	74	72	74	72	70	76
30	88	84	90	82	84	78	84	84	82	78	72

RAW TEST DATA (continued)

Var./S	34	78	72	86	94	60	68	25	75	15	44
31	74	72	74	72	72	70	78	66	78	72	76
32	76	78	66	76	86	66	86	76	74	74	76
33	78	72	73	68	68	74	72	60	71	68	76
34	51	44	45	37	40	47	44	30	43	40	49
35	72	72	72	72	72	71	72	72	72	72	74
36	43	43	43	43	43	42	43	43	43	43	46
37	78	77	80	73	73	68	73	68	74	67	73
38	48	47	50	42	42	36	42	36	43	35	42
39	24	16	13	13	18	13	19	4	12	7	15
40	51	45	43	43	47	43	48	36	42	39	45
41	6	9	5	13	7	4	9	6	8	6	8
42	40	47	38	56	43	36	47	40	45	40	45
43	20	16	19	27	18	25	18	21	22	30	24
44	40	34	39	52	37	49	37	42	44	57	47
45	46	43	43	46	42	42	44	38	43	42	46
46	15	19	8	17	25	31	25	22	28	20	23
47	39	42	33	41	47	52	47	45	50	43	45
48	54	109	96	57	80	71	81	72	76	75	93
49	33	58	52	34	45	41	45	41	43	43	51
50	30	34	11	52	13	23	15	22	57	31	30
51	42	44	33	53	34	39	35	38	56	43	42
52	78	40	78	55	79	77	67	77	0	40	1
53	56	36	56	44	56	55	50	55	16	36	17
54	42	45	44	43	46	47	44	45	41	41	39
55	44	44	44	44	44	44	44	42	42	42	42

Var./S	35	4	51	100	30	43	84	101	52	3	5
1	97	115	94	111	90	83	96	102	85	99	104
2	60	65	60	57	52	49	56	56	48	54	54
3	59	56	61	50	55	55	56	53	53	52	50
4	51	63	48	53	46	43	55	55	48	56	55
5	200	200	200	200	200	200	200	200	200	200	200
6	53	81	40	51	49	55	57	53	46	53	85
7	57	70	40	70	87	70	70	94	32	87	70
8	73	50	56	56	34	34	44	42	47	44	37
9	44	65	44	48	44	31	48	48	61	44	52
10	57	42	57	42	31	34	49	38	49	53	42
11	46	70	42	70	55	46	82	60	46	70	42
12	78	55	37	73	33	37	46	55	28	33	42
13	30	55	102	33	31	30	50	67	71	71	67
14	36	88	55	52	68	52	52	47	44	52	52
15	71	71	57	74	42	36	70	66	30	50	44

RAW TEST DATA (continued)

Var./S	35	4	51	100	30	43	84	101	52	3	5
16	81	66	67	75	45	40	79	68	21	60	41
17	57	64	51	53	64	46	49	49	32	54	56
18	56	56	51	46	68	52	48	46	40	52	52
19	78	78	42	60	63	36	36	69	30	60	42
20	57	60	48	54	57	42	48	51	39	48	54
21	42	54	42	54	57	42	42	45	30	66	42
22	60	54	54	54	66	30	54	42	36	48	60
23	62	72	72	42	62	54	48	30	30	42	84
24	45	54	42	42	102	84	54	54	48	54	48
25	9	14	3	11	13	3	10	8	8	8	8
26	40	52	26	45	49	26	42	38	38	38	38
27	26	40	22	44	45	10	17	36	17	14	7
28	69	74	68	77	76	68	67	73	66	66	66
29	70	70	66	84	72	66	66	72	66	66	66
30	68	84	68	94	88	70	68	76	74	70	66
31	66	74	72	70	74	72	72	74	66	66	66
32	72	74	68	62	78	60	62	66	60	60	60
33	80	68	74	60	71	74	64	60	66	71	68
34	53	40	47	30	43	47	35	30	37	43	40
35	71	72	72	60	72	72	64	61	70	61	64
36	42	43	43	25	43	43	31	27	40	27	31
37	71	65	68	62	73	72	61	65	65	66	60
38	40	33	36	29	42	41	28	33	33	34	27
39	4	13	3	13	17	6	4	9	10	5	11
40	36	43	36	43	46	38	36	40	41	37	42
41	1	6	8	7	5	2	3	5	6	6	3
42	29	40	45	43	38	31	34	38	40	40	34
43	28	18	17	19	19	15	13	14	23	13	26
44	54	37	35	39	39	32	28	30	45	28	50
45	42	39	40	35	42	39	32	33	39	35	37
46	12	11	8	3	15	7	13	13	9	0	1
47	37	36	33	29	39	32	37	37	34	27	28
48	34	70	71	67	54	92	86	48	26	79	35
49	24	40	41	39	33	50	47	30	20	44	25
50	50	29	22	41	44	12	9	18	14	9	5
51	52	42	38	48	49	33	32	36	34	32	30
52	61	65	73	75	0	31	65	70	56	10	2
53	47	49	53	54	16	32	49	52	45	21	17
54	40	42	41	42	34	37	41	39	33	31	25
55	41	40	40	38	38	38	36	36	36	33	31

BIBLIOGRAPHY

BIBLIOGRAPHY

Akutagawa, D., & Benoit, E. The effect of age and relative brightness on associative learning in children. Child Develop., 1959, 30, 229-238.

American Psychological Association, American Educational Research Association, and National Council on Measurements Used in Education, joint committee of. Technical recommendations for psychological tests and diagnostic techniques. Washington, D. C.: Amer. Psychol. Ass., 1954.

Anastasi, Anne. Psychological testing. (2nd ed.) New York: Macmillan, 1961.

Barnett, C. D. Stimulus generalization in normals and retardates on a visual-spatial task requiring a voluntary response. Unpublished doctoral dissertation, Peabody College, 1958.

Barnett, C. D., Ellis, N. R., & Pryer, Margaret W. Serial position effects in superior and retarded subjects. Psychol. Rep., 1960, 7, 111-113.

Binet, A., & Simon, T. Methodes nouvolles pour la diagnostic du niveau intellectus des anormaux. Annee psychol., 1905, 11, 191-244.

Bond, E. A. Tenth grade abilities and achievements. Teach. Coll. Contr. Educ., 1940, No. 813.

Bradway, Katherine P., Thompson, Clare W., & Cravens, R. B. Preschool IQ's after twenty-five years. *J. educ. Psychol.*, 1958, 49, 278-281.

Budoff, M., & Purseglove, E. M. PPVT performance of institutionalized mentally retarded adolescents. *Amer. J. ment. Defic.*, 1963, 67, 756-760.

Burke, H. R. Raven's Progressive Matrices: a review and critical evaluation. *J. genet. Psychol.*, 1958, 93, 199-228.

Burnett, A. Comparison of the Peabody Picture Vocabulary Test with the Wechsler-Bellevue and Stanford-Binet on educable mentally retarded children and adolescents. Owatonna, Minn.: Owatonna State School, 1964.

Buros, O. K. (Ed.) *The fifth mental measurement yearbook.* Highland Park, New Jersey: Gryphon Press, 1959.

Collman, R. D., & Newlyn, D. Changes in Terman-Merrill IQ's of mentally retarded children. *Amer. J. ment. Defic.*, 1958, 63, 307-311.

Cooley, W. W., & Lohnes, P. R. *Multivariate procedures for the behavioral sciences.* New York: Wiley, 1962.

Denny, M. R. Learning and performance. In H. A. Stevens & R. Heber (Eds.) *Mental retardation.* Chicago: Univer. of Chicago Press, 1964, Pp. 100-142.

Dunn, L. M. *Peabody picture vocabulary test.* Minneapolis: American Guidance Service, 1959.

Dunn, L. M. Expanded manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, 1965.

Eisman, Bernice S. Paired associate learning, generalization and retention as a function of intelligence. Amer. J. ment. Defic., 1958, 63, 481-489.

Ellis, N. R., Pryer, Margaret W., Distefano, M. K., Jr., & Pryer, R. S. Learning in mentally defective, normal, and superior subjects. Amer. J. ment. Defic., 1960, 64, 725-734.

French, J. L. Standardization of the Pictorial Intelligence Test. Paper read at annual convention of the American Psychological Association, St. Louis, August, 1962.

French, J. L. Pictorial test of intelligence. Boston: Houghton Mifflin, 1963.

Fruchter, B. Introduction to factor analysis. Princeton: Van Nostrand, 1954.

Goulet, L. R., & Barclay, A. The Vineland Social Maturity Scale: utility in assessment of Binet MA. Amer. J. ment. Defic., 1963, 67, 916-921.

Green, M. W., & Ervert, J. Normative data on Progressive Matrices (1947). J. consult. Psychol., 1955, 19, 139-142.

Jastak, J. Wide range achievement test. New York: Psychological Corp., 1946.

Jenson, G. III. A statistical analysis of the reliability and validity of the Pictorial Test of Intelligence and the Stanford-Binet Intelligence Scale. Unpublished masters thesis, Univer. of Missouri, 1964.

Kicklighter, R. *Comparison of PPVT and RSB test scores of educable mentally retarded children*. Atlanta: State Department of Education, 1964.

Kimbrell, D. L. Comparison of Peabody, WISC, and Academic Achievement scores among educable mental defectives. *Psychol*. *Rep*., 1960, 7, 502.

Kirk, S. A. Research in education. In H. S. Stevens & R. Heber (Eds.), *Mental retardation*. Chicago: Univer. of Chicago Press, 1964. Pp. 57-99.

Kirk, S. A., & Bateman, Barbara. Diagnosis and remediation of learning disabilities. *Except*. *Child*., 1962, 29, 73-78.

Kirk, S. A., Kass, Corinne, & Bateman, Barbara. The educability of psycholinguistic functions in retarded children. Urbana, Ill.: Institute for Research on Exceptional Children, Progress report, 1962.

Klaus, R. A., & Starke, Christiana. Experimental revision of the Peabody Picture Vocabulary Test as a predictor of first grade reading ability. Unpublished paper, Peabody College, 1964.

Lavitt, J. Comparison of the Peabody, Wechsler, Binet, and California tests of intellectual ability among 7th to 9th grade pupils. Unpublished paper, Westfield, Mass. Public Schools, 1963.

Lawson, J. R., & Avile, D. Comparison of Wide Range Achievement Tests and Gray Oral Reading Paragraphs reading scores of mentally retarded adults. *Percept. mot. Skills*, 1962, 14, 474.

Lindstrom, A. E. A comparison of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children. *Studies in Minnesota Education*, 1961, 131-132.

McCarthy, J. J., & Kirk, S. A. *The Illinois test of psycholinguistic abilities*. Madison, Wisc.: ITPA, Inc., 1961.

McCarthy, J. J., & Olson, J. L. *Validity studies on the Illinois test of psycholinguistic abilities*. Madison, Wisc.: Authors, 1964.

McCulloch, T. L., Reswick, J., & Roy, I. Studies of word learning in mental defectives: I. Effects of mental level and age. *Amer. J. ment. Defic.*, 1955, 60, 133-139.

Malpass, L. F., Brown, R., & Hake, D. The utility of the Progressive Matrices (1956 edition) with normal and retarded children. *J. clin. Psychol.*, 1960, 16, 350.

Martin, A. W., & Wiechers, J. E. Raven's Coloured Progressive Matrices and the Wechsler Intelligence Scale for Children. *J. consult. Psychol.*, 1954, 18, 143-144.

Mein, R. Use of the Peabody Picture Vocabulary Test with severely sub-normal patients. *Amer. J. ment. Defic.*, 1960, 65, 482-485.

Miller, M. B. Psychometric and clinical studies in mental deficiency, 1954-59; a selective review and critique. *Amer. J. ment. Defic.*, 1960, 65, 182-193.

Moss, J. W. An evaluation of the Peabody Picture Vocabulary Test with the PMA and 1937 Stanford-Binet. Unpublished paper, Univer. of Illinois, Institute for Research on Exceptional Children, 1962.

Mueller, M. W. Comparison of psycholinguistic patterns of gifted and retarded children. *J. school Psychol.*, in press.

Mueller, M., & Smith, J. O. The stability of language age modifications over time. *Amer. J. ment. Defic.*, 1954, 68, 537-539.

Mueller, M. W., & Weaver, S. J. Psycholinguistic abilities of institutionalized and non-institutionalized trainable mental retardates. *Amer. J. ment. Defic.*, 1964, 68, 775-783.

Osgood, C. E. Motivational dynamics of language behavior. *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1957.

Pascal, G. R., & Zak, M. Double alternation performance as a measure of educability in cerebral palsied children. *Amer. J. ment. Defic.*, 1956, 59, 658-665.

Pryer, R. S. Retroactive inhibition in normals and defectives as a function of temporal position of the interpolated task. *Amer. J. ment. Defic.*, 1960, 64, 1004-1011.

Raven, J. C. *Progressive Matrices: A perceptual test of intelligence, 1938, Sets A, B, C, D, and E.* London: Lewis, 1938.

Raven, J. C. *Progressive Matrices (1947).* London: H. K. Lewis and George G. Harrap, 1947.

Reed, J. C. The relationship between the primary mental abilities and reading achievement at given developmental levels. *Amer. Psychologist*, 1958, 13, 324.

Ringelheim, D. The relation of set to simple alternation problem solving in mental defectives. In R. L. Cromwell (Ed.) *Abstracts of Peabody studies in mental retardation*, 1960, 1, No. 26. (a)

Ringelheim, D. The relation of degree of set and CA to simple alternation problem solving in mental defectives. In R. L. Cromwell (Ed.) *Abstracts of Peabody studies in mental retardation*, 1960, 1, No. 27. (b)

Rohns, F. W., & Haworth, Mary R. The 1960 Stanford-Binet, WISC, and Goodenough tests with mentally retarded children. *Amer. J. ment. Defic.*, 1962, 66, 853-859.

Saxton, G. H., Blackman, L. S., & Tretakoff, M. I. Achievement measurement and academic grade placement in educable mental retardates. *Amer. J. ment. Defic.*, 1963, 67, 748-750.

Science Research Associates, Inc. *Primary Mental Abilities technical report*. Chicago: Authors, 1965.

Semmel, M. I., & Mueller, M. W. A factor analysis of the Illinois Test of Psycholinguistic Abilities with mentally retarded children. Unpublished paper, Peabody College, 1963.

Sievers, D. Development and standardization of a test of psycholinguistic growth on preschool children. Unpublished doctoral dissertation, Univer. of Illinois, 1955.

Silverstein, A. B. An evaluation of two short forms of the Stanford-Binet, Form L-M, for use with mentally retarded children. *Amer. J. ment. Defic.*, 1963(a), 67, 922-923.

Silverstein, A. B. Psychological testing practices in state institutions for the mentally retarded. *Amer. J. ment. Defic.*, 1963(b), 68, 440-445.

Sloan, W., & Berg, I. A. A comparison of two types of learning in mental defectives. *Amer. J. ment. Defic.*, 1957, 61, 556.

Sontag, L. W., Baker, C. T., & Nelson, V. L. Mental growth and personality development: a longitudinal study. *Monogr. Soc. Res. Child Develpm.*, 1958, 23, No. 2. Child Development Publications, Lafayette, Indiana.

Stacey, C. L., & Carleton, F. O. The relationship between Raven's Coloured Progressive Matrices and two tests of general intelligence. *J. clin. Psychol.*, 1955, 11, 84-85.

Stacey, C. L., & Gill, M. R. The relationship between Raven's Coloured Progressive Matrices and two tests of general intelligence for 172 subnormal adult subjects. *J. clin. Psychol.*, 1955, 11, 86-87.

Stolurow, L. M., & Pascal, G. R. Double alternation behavior in mental defectives. *Amer. Psychologist*, 1950, 5, 273-274.

Stott, D. H. Observations on retest discrepancy in mentally subnormal children. *Brit. J. educ. Psychol.*, 1960, 30, 211-219.

Tempero, H., & Ivanoff, J. *Effectiveness of the Peabody Picture Vocabulary Test with seventh-grade pupils*. Lincoln, Neb.: Univer. of Nebraska, Dept. Educ. Psychol. and Meas., 1960.

Terman, L. M., & Merrill, Maud A. *Stanford-Binet Intelligence Scale: Manual for the third revision, Form L-M.* Boston: Houghton Mifflin, 1960.

Throne, F. M., Schulman, J. L., & Kasper, J. C. Reliability and stability of the Wechsler Intelligence Scale for Children for a group of mentally retarded boys. *Amer. J. ment. Defic.*, 1962, 67, 455-457.

Thurstone, L. L., & Thurstone, Thelma G. *Examiners manual, PMA, Primary Mental Abilities for grades 2-4*. Chicago: Science Research Associates, 1962.

Tobias, J., & Gorelick, J. The validity of the PPVT as a measure of intelligence of retarded adults. *Train. Sch. Bull.*, 1961, 58, 92-98.

Walker, Helen M., & Lev, J. *Elementary statistical methods.* (Rev. Ed.) New York: Holt, Rinehart, Winston, 1958.

Wechsler, D. *Wechsler Intelligence Scale for Children.* New York: Psychol. Corp., 1949.

Weise, P. Current uses of Binet and Wechsler tests by school psychologists in California. *Calif. J. educ. Res.*, 1960, 11, 73-78.

Wilson, F. T. Learning of bright and dull children. *Teach. Coll. Contr. to Educ.*, 1928, 292.

Wolfensberger, W. The correlation between PPVT and achievement scores among retardates: A further study. *Amer. J. ment. Defic.*, 1962, 67, 450-451.

Woodcock, R. W. An experimental prognostic test for remedial readers. *J. educ. Psychol.*, 1958, 49, 23-27.

Woodcock, R. W. *Manual for the Prognostic Reading Test.* Oregon College of Education, 1960.

Woodcock, R. W. Modified Prognostic Reading Test. Unpublished material, Peabody College, 1964.

Wrightstone, J. W., Forlano, G., Lepkowski, J. R., Sontag, M., & Edelstein, J. D. A comparison of educational outcomes under single-track and two-track plans for educable mentally retarded children. U. S. Off. educ. Coop. Res. Br., Contract No. 6908, Project No. 144, 1959.